

23/5/4

DIALOG(R) File 351:Derwent WPI  
(c) 2001 Derwent Info Ltd. All rts. reserv.

008206371

WPI Acc No: 1990-093372/199013

Related WPI Acc No: 1990-101443

XRAM Acc No: C90-040856

Microbiological prodn. of human serum albumin - by culturing yeast contg.  
the HSA structural gene

Patent Assignee: RHONE-POULENC RORER SA (RHON ); RHONE-POULENC SANTE (RHON )

Inventor: FLEER R; FUKUHARA H; YEH P

Number of Countries: 004 Number of Patents: 006

Patent Family:

Patent No	Kind	Date	Applicat No	Kind	Date	Week
FR 2635115	A	19900209	FR 88635115	A	19880805	199013 B
ZA 8905821	A	19901031	ZA 895821	A	19890731	199048
JP 2276589	A	19901113	JP 89202715	A	19890804	199051
EP 361991	A3	19921014	EP 89402217	A	19890804	199340
DE 68929115	E	20000120	DE 629115	A	19890804	200011
			EP 89402217	A	19890804	
JP 3068139	B2	20000724	JP 89202715	A	19890804	200040

Priority Applications (No Type Date): FR 8810615 A 19880805; FR 88635115 A 19880805; FR 898897 A 19890703

Cited Patents: 6.Jnl.Ref; EP 241435; EP 258067; EP 301670; EP 308381; EP 319641; EP 79739

Patent Details:

Patent No	Kind	Lan Pg	Main IPC	Filing Notes
FR 2635115	A	71		
DE 68929115	E		C12P-021/02	Based on patent EP 361991
JP 3068139	B2	71	C12N-015/09	Previous Publ. patent JP 2276589

Abstract (Basic): FR 2635115 A

Prodn. of human serum albumin (HSA) is effected by culturing a yeast contg. an expression cassette comprising at least one selection marker together with DNA of the HSA structural gene fused to an expression promoter sequence and opt. a signal sequence for excretion of the HSA into the culture medium.

Also claimed are plasmids p7G19, p7G23, pKan-707 and pCXJ1, the yeast strain Kluyveromyces marxianus var. lactis MW98-8C, and the MW98-8C strain transformed with p7G19 or p7G23.

USE - HSA is used e.g. in the mfr. of blood extenders.

Dwg.o.0/30

Title Terms: MICROBIOLOGICAL; PRODUCE; HUMAN; SERUM; ALBUMIN; CULTURE; YEAST; CONTAIN; STRUCTURE; GENE

Derwent Class: B04; D16

International Patent Class (Main): C12N-015/09; C12P-021/02

International Patent Class (Additional): A61K-037/02; A61K-038/00; C07K-013/00; C07K-014/765; C07K-015/06; C12D-000/00; C12N-001/19; C12N-015/00; C12N-015/14; C12N-015/62; C12N-015/81; C12P-021/00; C12R-001/85; C12P-021/02; C12R-001-865; C12R-001-645

File Segment: CPI

+++++

(19) RÉPUBLIQUE FRANÇAISE  
 INSTITUT NATIONAL  
 DE LA PROPRIÉTÉ INDUSTRIELLE  
 PARIS

(11) N° de publication :  
 (à n'utiliser que pour les  
 commandes de reproduction)

**2 635 115**

(21) N° d'enregistrement national : **88 106 15**

(51) Int Cl<sup>s</sup> : C 12 P 21/00; C 12 N 15/00 // A 61 K 37/02;  
 (C 12 P 21/00, C 12 R 1:85).

(12)

**DEMANDE DE BREVET D'INVENTION**

A1

(22) Date de dépôt : 5 août 1988.

(71) Demandeur(s) : RHONE-POULENC SANTE, Société anonyme. — FR.

(30) Priorité :

(72) Inventeur(s) : Reinhard Fleer ; Hiroshi Fukuhara ; Patrice Yeh.

(43) Date de la mise à disposition du public de la demande : BOPI « Brevets » n° 6 du 9 février 1990.

(73) Titulaire(s) :

(60) Références à d'autres documents nationaux appartenants :

(74) Mandataire(s) : Rhône-Poulenc Interservices.

(54) Procédé de préparation de la sérum albumine humaine à partir d'une levure.

(57) Procédé microbiologique pour la préparation de la sérum albumine humaine par culture d'une levure, particulièrement du genre *Kluyveromyces*, modifiée par l'utilisation des techniques de l'ADN recombinant.

FR 2 635 115 - A1

D

Vente des fascicules à l'IMPRIMERIE NATIONALE, 27, rue de la Convention — 75732 PARIS CEDEX 15

La présente invention concerne un procédé microbiologique pour la préparation de la sérum albumine humaine (SAH) par culture d'une levure, particulièrement du genre Kluyveromyces, modifiée par l'utilisation des techniques de l'ADN recombinant.

5 La SAH est une protéine de 585 acides aminés, d'un poids moléculaire de 66.000 daltons, se présentant sous forme monomère globulaire et non glycosylée. Sa structure globulaire est maintenue par 17 ponts disulfures qui créent une série séquentielle de 9 boucles doubles<sup>1</sup>. Les gènes codant pour la SAH sont connus pour être hautement polymorphes et plus de 30 variants génétiques apparemment 10 différents ont été repérés par analyse électrophorétique dans des conditions variées<sup>2</sup>. Le gène de la SAH est coupé en 15 exons par 14 séquences introniques et comprend 16 961 nucléotides du site de "capping" supposé jusqu'au premier site d'addition de poly(A).

15 L'albumine humaine est synthétisée dans les hépatocytes du foie puis sécrétée dans le flux sanguin. C'est la protéine la plus abondante du sang avec une concentration d'environ 40 g/litre de sérum ; il y a donc environ 160 g d'albumine circulante dans le corps humain à tout moment. Le rôle le plus important de la SAH est de maintenir une osmolarité normale du flux sanguin. Elle présente également une capacité exceptionnelle de liaison pour diverses substances et joue un rôle aussi 20 bien dans le transport endogène de molécules hydrophobes (tels que les stéroïdes et les sels biliaires) que dans celui de différentes substances thérapeutiques qui peuvent ainsi être transportées à leurs sites d'action respectifs. De plus, la SAH a été récemment impliquée dans le catabolisme des prostaglandines.

25 La synthèse de SAH dans les hépatocytes conduit d'abord à un précurseur, la prépro-SAH, qui contient une séquence signal de 18 acides aminés dirigeant le polypeptide naissant dans la voie de la sécrétion. Cette séquence signal est coupée, probablement par un processus co-translationnel, avant que la protéine ne soit relarguée du réticulum endoplasmique. Ce premier clivage protéolytique

donne le précurseur pro-SAH qui contient encore à son extrémité N-terminale, un hexapeptide (Arg-Gly-Val-Phe-Arg-Arg) qui n'est pas normalement présent dans la forme mature de la SAH circulante.

Une convertase, probablement située dans le corps de Golgi, enlève  
 5 le propeptide au cours d'une deuxième étape de protéolyse en coupant le peptide lié  
 à la partie N-terminale de l'acide aspartique de la SAH<sup>3,4</sup>. Toutefois, la pro-  
 albumine humaine immature peut représenter la moitié de celle du flux sanguin chez  
 de rares individus hétérozygotes qui portent une mutation héréditaire dans la  
 séquence du propeptide, voire même de toute l'albumine chez certains homozygo-  
 10 tes extrêmement rares<sup>5</sup>. L'un de ces variants est désigné sous le nom de proalbumine  
 Christchurch et commence avec la séquence Arg-Gly-Val-Phe-Arg-Gln-Asp- dans  
 lequel la substitution Arg<sup>-1</sup> → Gln empêche la coupure du propeptide ainsi muté<sup>6</sup>.  
 D'autres variants appelés proalbumine Lille<sup>7</sup> et Takefu<sup>8</sup> ont respectivement les sé-  
 quences N-terminales suivantes : Arg-Gly-Val-Phe-Arg-His-Asp et Arg-Gly-Val-  
 15 Phe-Arg-Pro-Asp. Dans chacun des cas, la mutation observée concerne la paire des  
 acides aminés basiques Arg-Arg qui précède immédiatement l'albumine humaine  
 mature et provient du changement d'une seule base dans le codon arginine<sup>5</sup>.

La coupure après une paire d'acides aminés basiques est un point  
 essentiel dans la maturation d'autres protéines, y compris d'hormones peptidiques,  
 20 de neuropeptides et des protéines du plasma autres que la SAH<sup>9</sup>. Récemment, il  
 a été décrit que la convertase qui coupe la proalbumine en albumine dans le foie est  
 probablement très proche de la protéase yscF de la levure<sup>10</sup>. Cette thiol protéase,  
 calcium dépendante, est codée par le gène KEX2 de S. cerevisiae et est probable-  
 ment liée à la membrane de l'appareil de Golgi. Elle est connue pour être impliquée  
 25 dans la maturation de la phéromone appelé facteur alpha et dans celle de la toxine  
 "killer", puisque les mutants kex2 de S. cerevisiae sont incapables de couper les pro-  
 protéines respectives au niveau du couple d'acides aminés basiques Lys-Arg<sup>11</sup>. De  
 plus, on peut démontrer que l'enzyme de levure codée par le gène KEX2 reconnaît

correctement et coupe la proséquence normale (Arg-Arg) de l'albumine *in vitro*, mais ne coupe pas la proséquence mutée (Arg-Gln) de la proalbumine Christ-church<sup>4</sup>. Une convertase analogue à la protéase *yscF* a été récemment décrite comme codée par le gène *KEX1* de *K. lactis*<sup>11</sup>. Au vu de ce résultat, la capacité de ces levures pour maturer correctement l'albumine humaine recombinante et la sécréter, a été testée.

L'albumine représente 40 % du marché mondial des protéines plasmatiques. Son intérêt commercial réside dans le fait que ce produit est largement utilisé, par exemple, dans des solutés dits de remplissage pour compenser les pertes de sang au cours d'acte chirurgical, d'accident ou d'hémorragie, et à des doses pouvant atteindre plusieurs dizaines de grammes par jour et par individu.

Jusqu'à présent, la SAH est généralement produite par les techniques classiques de fractionnement du plasma provenant de dons de sang<sup>12</sup>, la consommation mondiale annuelle de plasma étant voisine de  $9,5 \times 10^6$  litres. Elle peut aussi être extraite du placenta humain selon la technique décrite par J. Liautaud et al.<sup>13</sup>

Le développement du génie génétique et des nouvelles techniques d'extraction et de purification a ouvert la possibilité d'obtenir des produits améliorés de plus haute pureté, de meilleure stabilité, sans contamination virale (par exemple hépatite B et SIDA) et à un prix de revient plus faible. Toutefois, aucun procédé assez performant basé sur le génie génétique n'est aujourd'hui connu permettant d'obtenir la SAH à l'échelle industrielle dans des conditions économiquement rentables.

Ceci est essentiellement dû à l'absence d'un couple hôte/vecteur performant qui permette la production importante d'une albumine sécrétée, correctement maturée, possédant une structure tertiaire conforme et ayant les propriétés physico-chimiques de l'albumine humaine native.

Même si l'utilisation de cultures de cellules de mammifères peut apparaître comme le choix idéal pour l'expression des protéines humaines, le prix de revient d'un tel procédé serait bien supérieur au prix de vente de l'albumine

pharmaceutique<sup>14</sup>. Comme ce produit est en général prescrit par dizaines de grammes et à un coût très bas, la production biotechnologique de la SAH n'apparaît réalisable qu'avec des systèmes de fermentation microbienne.

Jusqu'à une époque récente, la grande majorité des expériences de génie génétique ont utilisé *E. coli* comme organisme hôte pour la production par voie microbienne de protéines hétérologues économiquement importantes. Même si ce type de procédé est apparu satisfaisant pour un bon nombre de protéines hétérologues, tous les efforts pour produire la SAH dans des conditions économiquement acceptables en utilisant cet organisme n'ont été que partiellement couronnés de succès.

L'un des problèmes majeurs associé à l'utilisation de *E. coli* réside dans le fait que cette bactérie est, dans la majorité des cas, incapable d'excréter les protéines hétérologues dans le milieu de culture et que celles-ci s'accumulent dans le cytoplasme, le périplasme et les membranes. La protéine recherchée doit donc être séparée des constituants cellulaires, ce qui conduit à des procédés complexes et coûteux. De plus, *E. coli* produit des endotoxines et des pyrogènes qui peuvent contaminer les protéines produites. Pour cette raison, la plus grande attention doit être portée pendant la purification du taux résiduel d'endotoxines dans le produit final, spécialement s'il est utilisé comme médicament.

Les protéines sécrétées par leur hôte naturel perdent souvent la propriété de se replier correctement si elles sont synthétisées et accumulées dans le cytoplasme. En général, la sécrétion est requise pour permettre la formation des ponts disulfures comme ceux présents dans la SAH. Le fait que la SAH produite dans *E. coli* ne soit pas sécrétée conduit ainsi à sa précipitation intracellulaire sous forme insoluble<sup>15</sup>. En conséquence, après extraction à partir des cellules bactériennes, celle-ci doit être dénaturée puis renaturée *in vitro*.

De plus, *E. coli* ne possède pas la machinerie cellulaire permettant de réaliser les modifications post-transcriptionnelles et post-traductionnelles spéci-

ques des organismes eucaryotes, y compris l'homme. En ce qui concerne la SAH, il semble difficile d'exprimer à un taux satisfaisant dans *E. coli* le précurseur naturel de SAH, c'est-à-dire la prépro-SAH. D'autre part, des proalbumines artificielles dans lesquelles le gène de structure de la SAH a été fusionné à des signaux de sécrétion dérivés des gènes bactériens *pac* ou *ompA* sont en général maturés de façon incomplète<sup>15</sup>. L'expression de la SAH sans la présequence, c'est-à-dire sous forme Met-SAH, montre que le résidu méthionine N-terminal n'est pas excisé par la méthionine aminopeptidase de *E. coli* (MAP)<sup>15,16</sup>. En conséquence, la SAH de *E. coli* ne peut pas être obtenue sous forme mature *in vivo* et doit être modifiée *in vitro*.

10 par coupure trypsique, afin d'exciser une séquence leader dérivée d'un phage et ce après dénaturation et renaturation *in vitro*<sup>17</sup>.

En ce qui concerne les autres hôtes bactériens, des travaux concernant la sécrétion de SAH dans *Bacillus subtilis* ont été récemment publiés<sup>18</sup>. Bien que cette étude indique qu'un organisme procaryote tel que *B. subtilis* possède le potentiel pour sécréter des protéines humaines de haut poids moléculaire comme la SAH, elle montre aussi que l'excrétion dans le milieu de culture ne peut être obtenue qu'en utilisant des protoplastes de *B. subtilis*, c'est-à-dire des cellules dont la paroi cellulaire normale a été digérée par traitement enzymatique. De plus, le pourcentage de la SAH modifiée est inversement proportionnel au taux de protéine produite:

15 à des niveaux d'expression élevés, la plus grande partie de SAH recombinante reste sous forme immature<sup>19</sup>. En outre, aucune donnée concernant les propriétés physico-chimiques de la SAH recombinante de *B. subtilis* n'a été mentionnée. D'autre part, le niveau d'excrétion de SAH obtenu par ce microorganisme est très faible au vue des quantités de SAH détectées par utilisation de méthodes immunologiques dans

20 les surnageants de culture<sup>19</sup>.

25

L'utilisation de systèmes microbiens eucaryotes tels que les levures ou les champignons représente une alternative intéressante d'utilisation d'hôtes bactériens pour la préparation de protéines hétérologues. En effet, ces organismes pos-

sèdent toutes les caractéristiques de structure et d'organisation cellulaire des organismes eucaryotes plus complexes, telles que les cellules de mammifère. En particulier, les levures sont capables de réaliser les modifications post-transcriptionnelles et post-transductionnelles importantes pour l'activité de nombreuses protéines. En outre, les levures sont bien connues à l'échelle industrielle : elles peuvent être cultivées à haute densité cellulaire, elles ne sont pas pathogènes, ne produisent pas d'endotoxines et sont utilisées dans l'industrie alimentaire depuis très longtemps. Enfin, contrairement aux cellules de mammifères, les manipulations génétiques sur les levures sont faciles à réaliser et les données de la génétique classique et moléculaire sont nombreuses.

Le terme levure est fréquemment employé pour Saccharomyces cerevisiae, ou levure du boulanger, qui constitue l'une des espèces les plus communes et les mieux connues. Il est entendu dans ce qui suit que le terme levure s'applique à d'autres genres et n'est pas restreint à l'espèce S. cerevisiae.

La sécrétion de la SAH dirigée par son propre peptide signal sous le contrôle du promoteur de la chélatine a pu être mise en évidence dans S. cerevisiae à des taux maximaux d'environ 1 % des protéines totales<sup>20</sup>. Toutefois, le matériel reconnu par les anticorps anti-SAH décrit dans cette étude reste associé à la cellule et n'est pas exporté dans le surnageant de culture. De plus, aucune caractérisation détaillée de la protéine recombinante n'est rapportée.

La production de la SAH dans la levure de brasserie en utilisant un procédé post-fermentaire lors de la fabrication de la bière a également été mentionnée<sup>21</sup>. De nouveau aucune donnée quantitative ou qualitative concernant le produit obtenu n'a été décrite. De plus, ce procédé conduit à l'expression de Met-SA<sub>H</sub>, c'est-à-dire d'un allèle de la SAH démarrant par l'acide aminé méthionine juste en amont de la séquence de l'albumine mature. L'absence d'une séquence signal empêche la sécrétion et la maturation de la SAH recombinante et provoque l'accumulation d'une albumine intracellulaire dont la structure tertiaire n'a pas été caractérisée. De plus, l'absence de méthionine N-terminale dans le produit final n'a

pas été montrée.

La présente invention décrit l'obtention de souches de levures modifiées par génie génétique pouvant être cultivées en masse et capable de produire et d'excréter efficacement dans le milieu de culture la SAH dans sa conformation naturelle.

Un système d'expression préféré implique des levures du genre Kluyveromyces comme hôte et utilise des vecteurs dérivés du plasmide naturel pKD1 de K. marxianus var. drosophilicarum.

Les levures du genre Kluyveromyces et en particulier K. marxianus (incluant les variétés lactis et marxianus qui, dans ce qui suit, vont être appelées K. lactis et K. fragilis) sont des organismes importants et d'un intérêt commercial considérable dans l'industrie biotechnologique. K. lactis et K. fragilis sont utilisées, par exemple, pour la production commerciale de l'enzyme lactase ( $\beta$ -galactosidase). Ces levures sont capables de croître sur lactoserum, un sous-produit majeur de l'industrie laitière. Plusieurs souches de Kluyveromyces sont utilisées pour la production à grande échelle de "protéines d'origines unicellulaire" (P.O.U.) qui jouent un rôle important dans l'alimentation du bétail. Enfin, les organismes du genre Kluyveromyces sont mentionnés sur la liste G.R.A.S. (Generally Recognized As Safe), ce qui représente un facteur important en vue de la production de produits de qualité pharmaceutique.

Contrairement aux progrès importants réalisés dans le domaine de la génétique moléculaire de S. cerevisiae, les techniques de manipulation génétique n'ont été développées que récemment chez Kluyveromyces. Trois types de vecteurs de clonage ont été décrits chez cet organisme :

i) des vecteurs intégratifs contenant des séquences homologues de régions du génome de Kluyveromyces et qui, après introduction dans les cellules, s'intègrent dans les chromosomes de Kluyveromyces par recombinaison *in vivo*. L'intégration, bien qu'étant un événement très rare nécessitant la présence d'un marqueur de sélection efficace, est obtenue lorsque ces vecteurs ne contiennent pas de séquence permettant une réplication autonome dans la cellule. L'avantage de ce

système est la stabilité des souches transformées, c'est-à-dire le fait qu'elles puissent être cultivées dans un support nutritif normal sans le besoin d'une pression de sélection pour le maintien des séquences intégrées. L'inconvénient est toutefois que les gènes intégrés sont seulement présents à une ou, au mieux, à un petit nombre de copies par cellule. Le faible dosage des gènes résulte souvent en un faible niveau de production d'une protéine hétérologue.

5           ii) des vecteurs répliquatifs contenant des séquences de réPLICATION au-tonome (ARS) dérivées de l'ADN génomique de Kluyveromyces sp.<sup>22,23,24</sup> (KARS). De tels vecteurs ne semblent pas très intéressants puisque leur ségrégation lors de  
10          la mitose est très peu homogène, ce qui résulte en leur perte à un taux élevé, même lorsque celles-ci sont cultivées sous pression de sélection.

15          iii) des vecteurs répliquatifs contenant des ARS dérivés de plasmides naturels de levure, soit du plasmide linéaire " killer " pGKI-1 isolé de K. lactis<sup>25,26</sup>, soit du plasmide circulaire pKD1 isolé de K. drosophilicola<sup>27</sup>. Les vecteurs conte-nant des séquences ARS dérivées du plasmide linéaire " killer " n'ont pas d'utilité pratique en vue de la production en masse de protéines hétérologues puisqu' un milieu nutritif particulier est nécessaire à leur maintien et puisqu'ils sont perdues à un taux de 40-99 % dans une population donnée après seulement 15 générations, même sous pression de sélection<sup>26</sup>. Le système vecteur le plus efficace pour la trans-  
20          formation du genre Kluyveromyces décrit jusqu'alors est dérivé du plasmide endogène pKD1: des constructions contenant la séquence entière de pKD1 peuvent être transformées à haute fréquence, sont présentes dans la cellule à 70-100 copies, et, ce qui est le plus important, ont une stabilité assez élevée en conditions non sélectives. Néanmoins, l'efficacité des vecteurs décrits dans la demande EP 0 241 435  
25          demeure limitée à des applications de recherche dans la mesure où des fermenta-tions à l'échelle industrielle requièrent une stabilité du plasmide pendant au moins 40 générations. Même le vecteur le plus performant (P3) décrit dans la demande EP 0 241 435 est perdu par environ 70 % des cellules d'une population donnée après seulement six générations en milieu non sélectif<sup>27</sup>. Bien qu'il soit techniquement  
30          possible de maintenir une pression de sélection à grande échelle, l'utilisation d'un

milieu sélectif résulte souvent en une densité cellulaire fortement diminuée, et nécessite le recours à des souches bien définies ce qui rend cette approche moins attrayante et plus onéreuse. De plus, la demande EP 0 241 435 ne contient pas un seul exemple d'expression de protéine hétérologue d'intérêt commercial. En fait, le 5 seul gène "hétérologue" dont l'expression est démontrée à partir de vecteurs dérivés de pKD1 est le gène URA3 de la levure S. cerevisiae. Il reste donc à prouver que l'introduction dans pKD1 d'un gène non issu d'une levure, par exemple un gène d'origine mammifère, et sa surexpression dans Kluyveromyces n'entraînent pas une instabilité accrue du plasmide.

10 Un objet de la présente invention est l'obtention de nouveaux vecteurs d'expression capables de transformer des levures du genre Kluyveromyces et possédant des caractéristiques de stabilité nettement supérieures à celles mentionnées dans la demande EP 0 241 435 . Il sera démontré que les nouvelles constructions sont maintenues à haut nombre de copies dans 85-90 % des cellules après 50 générations de croissance en milieu non sélectif. Il est ainsi possible de produire des protéines hétérologues à partir de vecteurs multicopies stables en utilisant des souches industrielles du genre Kluyveromyces, qui ont été utilisées pendant de nombreuses années en raison de leurs propriétés optimales de croissance et à des densités cellulaires élevées.

15 20 La stabilité élevée des vecteurs décrits dans la présente invention a été obtenue en exploitant pleinement les caractéristiques du plasmide pKD1. Les vecteurs dérivés de pKD1 diffèrent de tous les autres vecteurs connus de Kluyveromyces en ce qu'ils sont caractérisés par un système spécialisé de réPLICATION responsable de leur maintien stable à haut nombre de copies. En plus d'une origine de réPLICATION (ARS), ce système comprend deux séquences répétées-inversées, chacune de 346 nucléotides de long, et trois phases ouvertes de lecture (gènes A, B et C) qui font partie intégrante du plasmide<sup>30</sup>. Par analogie avec le système plus étudié du plasmide 2u de S. cerevisiae qui est structurellement proche, les protéines codées par deux de ces gènes (B et C) sont vraisemblablement impliquées dans la partition 25 30 du plasmide lors de la mitose et pourraient jouer un rôle dans la régulation négative

du gène **A** codant pour une recombinase à site spécifique (FLP)<sup>23</sup>. Il a été montré que la recombinaison FLP-dépendante entre les séquences répétées inversées de l'ADN de 2  $\mu$  fait passer le plasmide d'un mode de réPLICATION régulier (un doublement des plasmides 2  $\mu$  par division cellulaire) à un mode de réPLICATION du type "rolling circle" (amplification du nombre de copies à environ 50 copies par cellule)<sup>23</sup>. Cette altération du mode normal de réPLICATION est induite aussitôt que le nombre de copies du plasmide devient trop bas pour permettre la production de quantités suffisantes des produits des gènes **B** et **C** qui agissent comme réPRESSEURS du gène **A** codant pour la recombinase FLP. Par ce mécanisme, le nombre de copies du 2  $\mu$  (et très probablement de plasmides structurellement semblables tels que pKD1) est maintenu à des niveaux élevés de façon autorégulée et est indépendant de la présence d'un marqueur de sélection.

Les vecteurs précédemment publiés dans la demande EP 0 241 435 soit contiennent seulement une partie de pKD1 (A15) soit un gène **A** interrompu (P1 et P3 pour lesquels le site de clonage PstI est situé dans la séquence codante du gène **A**<sup>27</sup>), détruisant ainsi le système de réPLICATION autorégulé qui est une des caractéristiques du plasmide résistant pKD1. Au contraire, les constructions plasmidiques dérivées de pKD1 décrites dans la présente invention respectent l'intégralité fonctionnelle de toutes les phases ouvertes importantes de pKD1. En conséquence, la stabilité des plasmides décrits ci-après est considérablement renforcée par rapport aux plasmides P1 et P3 et le nombre de copies des nouveaux vecteurs est maintenu à un niveau élevé dans toute la population cellulaire.

Dans la présente invention, le terme SAH sera utilisé pour désigner toute albumine du sérum d'origine humaine ou pour toute protéine isolée à partir de cellules de levure et présentant la même séquence d'acides aminés, la même structure tertiaire et les mêmes propriétés physico-chimiques que la SAH naturelle d'origine humaine.

La présente invention concerne la création par génie génétique de nouveaux vecteurs permettant la production de SAH sous une forme facilement purifiable en utilisant des cellules de levure comme hôte eucaryote microbien. Le

procédé de production est caractérisé par une excrétion efficace de SAH mature, native dans le milieu de culture des dites levures, ce qui facilite considérablement la purification de la protéine recombinante. La production de SAH est obtenue en cultivant des levures transformées par un plasmide recombinant comprenant une 5 région de démarrage de la transcription et de la traduction fonctionnant efficacement dans l'hôte et une phase ouverte de lecture codant pour la SAH précédée d'un peptide signal dirigeant la protéine recombinante dans la voie de sécrétion desdites levures.

10 Selon la présente invention, les nouvelles constructions incluent des cassettes d'expression contenant le gène structural de la SAH. Les constructions sont habituellement préparées de façon séquentielle en employant les vecteurs appropriés, jusqu'à ce que les éléments indispensables à l'expression, la sécrétion, ou au maintien du plasmide soient combinés pour former un vecteur qui peut alors être introduit dans le (ou les) hôte(s) purifié(s) de façon à exprimer et excréter la protéine 15 désirée. Les hôtes employés sont d'origine microbiologique eucaryote, en particulier des levures, plus particulièrement des genres Saccharomyces ou Kluyveromyces et de préférence de l'espèce Kluyveromyces marxianus incluant toutes ses variétés, en particulier K. marxianus var. lactis. Ainsi, selon la présente invention, les constructions qui sont préparées et décrites ont valeur d'exemples dans les organismes 20 eucaryotes d'origine microbiologique, mais sont destinées de préférence à Kluyveromyces.

Les hôtes particuliers à employer seront préférentiellement des souches industrielles stables, atteignant des hautes densités cellulaires dans les milieux appropriés et caractérisées par un niveau élevé de production.

25 La séquence codante pour la SAH peut être obtenue de diverses façons, la plus simple consistant à isoler des ARN messagers de foie humain et à synthétiser leurs copies sous forme d'ADN complémentaire (ADNc). Les séquences clonées peuvent ensuite être modifiées par différentes méthodes telles que la mutagénèse *in vitro*, l'elongation d'amorces, la restriction, l'insertion d'adaptateurs 30 ou la ligature avec des oligonucléotides de raccord. La séquence codante peut être

adaptée, par exemple, à l'usage des codons chez la levure pour optimiser l'efficacité de la synthèse protéique (traduction).

A l'extrémité N-terminale de la séquence protéique, un peptide signal peut être introduit de façon à diriger la protéine en cours de synthèse dans la voie de sécrétion de la cellule hôte. Cette séquence-signal peut correspondre à l'extension N-terminale naturelle de l'albumine ou peut être obtenue à partir de gènes de levure tels que ceux codant pour la phéromone alpha ou la toxine " killer ".

De plus, une proséquence, codant pour une autre extension peptidique, peut être intercalée entre la séquence-signal de sécrétion et la séquence codante de l'albumine mature. Cette séquence est normalement jointe à la séquence codante par l'intermédiaire d'un site de coupure d'une protéase spécifique impliquant en général au moins deux acides aminés de type basique, de préférence Lys-Arg ou Arg-Arg.

La cassette d'expression comprendra une région de démarrage de la transcription et de la traduction jointe à l'extrémité 5' terminale de la séquence codante, de façon à diriger et à réguler la transcription et la traduction de la dite séquence. Le choix de ces régions promotrices peut varier en fonction de l'hôte utilisé en particulier. Généralement, ces séquences sont choisies parmi les promoteurs dérivés des gènes de levure. D'un intérêt particulier sont certaines régions issues de gènes glycolytiques de levure du type Saccharomyces ou Kluyveromyces, telles que les gènes codant pour la phosphoglycérate kinase (PGK), la glyceraldehyde-3-phosphate deshydrogénase, la lactase, l'énoïlase, l'alcool deshydrogénase, etc. Ces régions de contrôle peuvent être modifiées, par exemple par mutagénèse in vitro, par l'introduction d'éléments additionnels de contrôle ou de séquences synthétiques ou par des délétions. Par exemple, des éléments de régulation de la transcription, telles que les régions dites UAS ("upstream activating sequences") provenant d'un autre promoteur comme celui du gène GAL10 de S. cerevisiae ou LAC4 de K. lactis peuvent être utilisés pour construire des promoteurs hybrides qui permettent de dissocier la phase de croissance d'une culture de levures de la phase d'expression d'un gène héterologique en fonction de la source carbonique choisi. Une région permettant une ter-

minaison de la traduction et de la transcription fonctionnelle dans l'hôte envisagé sera positionnée en 3' de la séquence codante.

La cassette d'expression ainsi construite sera fusionnée à un ou plusieurs marqueurs permettant de sélectionner l'hôte transformé. Les marqueurs préférés dans la levure sont des marqueurs dominants, c'est-à-dire conférant une résistance à des antibiotiques comme la génétidine (G418) ou tout autre composé toxique tels que des ions cuivre, dans la mesure où ceux-ci peuvent être utilisés sans spécificité d'hôte particulière. Ces gènes de résistance seront placés sous le contrôle de signaux appropriés de transcription et de traduction dans l'hôte considéré.

L'ensemble constitué par la cassette d'expression et par le marqueur de sélection peut être utilisé soit pour transformer l'hôte directement, soit peut être inséré sur un plasmide vecteur. Dans le cas d'une transformation directe, des séquences homologues à des régions présentes sur un chromosome ou sur un plasmide résistant sont fusionnées à cet ensemble. Les dites séquences seront situées de chaque côté de la cassette d'expression de façon à induire une insertion au site homologue par recombinaison *in vivo*. Si la cassette d'expression est insérée sur un plasmide vecteur, elle devra être combinée à un système de réPLICATION fonctionnel dans l'hôte considéré. Un système de réPLICATION préféré pour *Kluyveromyces* est dérivé du plasmide pKD1, initialement isolé de *K. drosophilaeum*<sup>29,30</sup>. Un système préféré de réPLICATION pour *Saccharomyces* est dérivé du plasmide de levure 2 $\mu$ . Le plasmide d'expression peut contenir tout ou partie desdits systèmes de réPLICATION ou peut combiner des éléments dérivés de pKD1 aussi bien que du plasmide 2 $\mu$ . Les constructions préférées sont celles qui contiennent la séquence entière du plasmide pKD1 lorsque l'expression dans *Kluyveromyces* est recherchée. Plus particulièrement, les constructions préférées sont celles où le site d'insertion des séquences étrangères dans pKD1 est localisé dans une région de 197 pb située entre le site SacI (position 4714 de la forme B de pKD1<sup>30</sup>) et MstII (position 154 de la forme B de pKD1<sup>30</sup>) ou à l'un de ces deux sites.

Pour plus de commodité, le plasmide peut être un vecteur navette : comme tel, il peut être transféré à un hôte bactérien tel que *E. coli*, où il peut être

manipulé plus facilement que dans la levure à chaque étape de construction. Dans ce cas, une origine de réPLICATION et un marqueur de sélection fonctionnant dans l'hôte bactérien sont requises. Il est également possible de positionner des sites de restriction qui sont uniques sur le vecteur d'expression de façon qu'ils entourent les 5 séquences bactériennes. Ceci permet de supprimer l'origine de réPLICATION bactérienne par coupure et religature du vecteur avant transformation des cellules de levure , ce qui peut résulter en une augmentation du nombre de copies et en une stabilité accrue du plasmide. Des sites pratiques tels que 5'-GGCCNNNNNGGCC- 10 3' (SfiI) ou 5'-GCGGCCGC-3' (NotI) peuvent par exemple être utilisés dans la mesure où ils sont extrêmement rares chez la levure et généralement absents sur un plasmide d'expression. Ces sites peuvent être introduits sur le vecteur par mutagénèse dirigée par oligonucléotide ou par l'addition d'oligonucléotides de raccord spécifiques.

Une fois terminée la construction du vecteur d'expression, celui-ci 15 sera introduit dans l'hôte désiré. Divers protocoles de transformation ont été décrits dans le cas des levures<sup>31</sup>. Les transformants peuvent alors être mis en culture dans un milieu approprié en vue de la production du produit recherché.

Si le produit recherché est la SAH excretée, il peut être purifié de diverses manières à partir du surnageant de culture. Il peut être cryoprécipité, extrait 20 par chromatographie d'affinité, par électrophorèse ou par d'autres techniques conventionnelles. La récupération de la protéine sécrétée peut être faite à partir d'une culture en " batch " ou en continu.

Les exemples suivants , donnés à titre non limitatif, montrent comment l'invention peut être mise en pratique.

#### Techniques générales de clonage

Les méthodes classiques de biologie moléculaire telles que la centrifugation d'ADN plasmidique en gradient de chlorure de césium - bromure d'éthidium, la

digestion par les enzymes de restriction, l'électrophorèse sur gel, l'électroélution des fragments d'ADN à partir de gels d'agarose, la transformation dans E. coli, etc, sont décrites dans la littérature<sup>32,33</sup>.

Les enzymes de restriction ont été fournies par New England Biolabs  
5 (Biolabs), Bethesda Research Laboratories (BRL) ou Amersham.

Pour les ligatures, les fragments d'ADN sont séparés selon leur taille sur des gels d'agarose 0,7% ou acrylamide 8%, purifiés par électroélution, extraits au phénol, précipités à l'éthanol puis incubés dans un tampon Tris-HCl 50 mM, MgCl<sub>2</sub> 10 mM, dithiothreitol 10 mM, ATP 2 mM, pH 7,4, en présence d'ADN ligase  
10 du phage T4 (Biolabs).

Si nécessaire, les fragments d'ADN ayant des extrémités 5' proéminentes sont déphosphorylés par un traitement à la phosphatase alcaline d'intestin de veau (CIP, Pharmacia) à 37°C pendant 30 minutes dans le tampon suivant: glycine 100 mM, MgCl<sub>2</sub> 1 mM, ZnCl<sub>2</sub> 1 mM, pH 10,5. La même technique est utilisée pour la déphosphorylation des extrémités 3' proéminentes ou franches, mais le traitement est de 15 minutes à 37°C puis de 15 minutes à 56°C. L'enzyme est inactivée par chauffage du mélange réactionnel à 68°C pendant 15 minutes en présence de SDS 1 % et de NaCl 100 mM suivi d'une extraction au phénol-chloroforme et d'une précipitation à l'éthanol.  
15

Le remplissage des extrémités 5' proéminentes est effectué par le fragment de Klenow de l'ADN polymérase I d'E. coli (Biolabs). La réaction est effectuée à température ambiante pendant 30 minutes dans un tampon Tris-HCl 50 mM, dNTPs 0,4 mM, MgSO<sub>4</sub> 10 mM, dithiotréitol 0,1 mM, BSA (Bovine Serum Albumine) 50 ug/ml, pH 7,2. Le remplissage des extrémités 3' proéminentes est effectué en présence de l'ADN polymérase du phage T4 (Biolabs) selon les recommandations du fabricant. La digestion des extrémités proéminentes est effectuée par traitement limité à la nucléase S1 (BRL) selon les recommandations du fabricant.  
20  
25

La mutagénèse dirigée in vitro par oligodésoxynucléotides est effectuée selon la méthode développée par Eckstein et coll.<sup>34</sup> en utilisant le kit distribué  
30

par Amersham.

Les ADNs ligaturés sont utilisés pour transformer la souche rendue compétente : *E. coli* MC1060 ( [*lac*IPOZYA], X74, *galU*, *galK*, *strA*\*) ou TG1 ( [*lac*pro A,B], *supE*, *thi*, *hsdD5/FtraD36*, *proA*<sup>+</sup>*B*<sup>+</sup>, *lacI*<sup>q</sup>, *lacZ* M15). L'ADN plasmidique est purifié à partir des transformants résistants à l'ampicilline ou à la tétracycline selon les cas. L'extraction de l'ADN plasmidique est faite selon le protocole décrit dans Maniatis et coll.<sup>32</sup> qui dérive de la méthode de lyse alcaline décrite par Birnboim et Doly<sup>33</sup>. Pour l'analyse rapide des plasmides, des lysats bactériens sont préparés selon la méthode de Holmes et Quigley<sup>34</sup> et sont analysés par électrophorèse sur gels d'agarose sans purification préalable. Après analyse par les endonucléases de restriction, les plasmides recombinants présentant la structure désirée sont préparés à plus grande échelle selon le protocole de lyse alcaline<sup>32</sup> à partir de cultures de 0,5 à 1 litre et purifiés par une centrifugation isopycnique en chlorure de césium.

Les transformations de *K. lactis* avec de l'ADN étranger ainsi que les purifications d'ADN plasmidiques de *K. lactis* sont décrites dans le texte.

**Exemple 1 : CLONAGE D'UNE CASSETTE CONTENANT LE GENE STRUCTURAL DE LA prépro-SAH**

**E.1.1 Isolement d'un clone d'ADN complémentaire (ADNc) codant pour la SAH**

La construction de plasmides recombinants permettant l'expression de la SAH dans *E. coli* a été décrite en détails dans la précédente demande de brevet EP 0198745<sup>16</sup>. En résumé, l'ADNc a été obtenu à partir d'ARNm poly-A isolés de foie humain selon la technique au thiocyanate de guanidium. L'analyse de la séquence d'ADN a permis l'isolement de trois clones (pT1B11, pAA38 et p6D8 figure 1) qui contenaient des fragments chevauchant du gène structural de l'albumine et présentaient des sites de restriction communs dans les régions chevauchantes. Ces sites ont été utilisés dans la reconstruction d'un clone d'ADNc de la SAH (figure 1) qui est complet sauf pour sa séquence "pré-pro".

### E.1.2 Synthèse de la séquence signal de la SAH

Dans le plasmide pXL276 (figure 2), dont la construction est décrite en détail dans la demande de brevet EP 0198745, la séquence codante de la SAH mature a été fusionnée en phase au codon ATG du site de fixation des ribosomes (RBS) du gène cII du bactériophage lambda. Ceci a généré un site NdeI immédiatement en amont du codon de démarrage de la traduction du gène de la SAH. pXL276 est utilisé pour reconstituer la région "pré-pro" de la SAH, laquelle se trouvait être tronquée dans l'ADNc juste en amont du site TaqI au niveau de l'acide aminé "-1". La reconstitution est effectuée en insérant un fragment d'ADN correspondant à la séquence signal de la SAH sous forme de quatre oligodéoxynucléotides de 33-36 bases de long (Sq32, Sq34, Sq35 et Sq38 ; figure 2) et un fragment EcoRI-NdeI de 120 pb provenant de pXL276 portant des signaux d'expression de gènes d'E. coli entre les sites EcoRI et AccI du plasmide pUC8. Le site TaqI est ainsi reconstitué à l'une des extrémités de l'insertion dans ce plasmide appelé pXL290 (figure 2). Le petit fragment HindIII-TaqI portant le RBS et la séquence prépro en amont du site TaqI est ligaturé avec le fragment TaqI-PstI du clone d'ADNc de la SAH (plasmide p1B11, contenant l'extrémité 5' de la SAH) entre les sites HindIII et PstI de pUC8 pour donner le plasmide pXL299 (figure 3). Le plasmide pXL322 est construit en ligaturant le fragment HindIII-PvuII de pXL299, portant le RBS (site de fixation de ribosome) et la séquence prépro -SAH jusqu'au site PvuII, avec le fragment PvuII-EcoRI de pXL276, portant le réplicon et l'extrémité 3' de la séquence codante de la SAH, et avec le fragment EcoRI-HindIII de pXL276 contenant le promoteur (figure 4).

### E.1.3 Crédation d'un site HindIII en amont du codon de démarrage de la traduction de la SAH.

Dans le but d'obtenir une cassette prépro- SAH pouvant être facilement intégrée dans des vecteurs d'expression, le site NdeI du plasmide pXL322, décrit ci-dessus, est changé en un site HindIII par mutagénèse dirigée par oligodéoxynucléotides. A cette fin, le fragment HindIII-BglII de pXL322 contenant

l'extrême 5' du gène de la prépro-SAH est sous-cloné dans M13mp18 et mutagénisé par hybridation de l'oligodéoxynucléotide synthétique 5'-ATCTAAGGAAA TACAAGCTTATGAAGTGCGGT-3' à la matrice simple brin (les séquences soulignées et en caractère gras représentent respectivement le site HindIII et le site de démarrage de la traduction). On obtient ainsi le plasmide pXL855 (figure 5) dont la séquence nucléotidique de la région mutagénisée est ensuite vérifiée. La séquence codant pour la SAH complète est reconstituée en insérant le fragment HindIII-PvuII du phage mutagénisé et le fragment PvuII-HindIII, contenant l'extrême 3' du gène structural de la SAH, dans le site HindIII de pUC8 pour donner le plasmide pXL869 (figure 6). Ce plasmide contient donc un fragment HindIII de 1,87 kb contenant le gène structural de la prépro-SAH complet ainsi qu'une région de 61 pb non-traduite à son extrémité 3'. La séquence complète du fragment HindIII ainsi que la séquence en acides aminés de la SAH recombinante sont décrites sur la figure 7. Une cassette HindIII contenant le gène structural de Met-SAH sans son signal de sécrétion est construite en suivant le même protocole sauf qu'un dérivé M13 du plasmide pXL276 est mutagénisé à l'aide de l'oligodéoxynucléotide synthétique 5'-ATCTAAGGAAA TACAAGCTTATGGATGCACACAAG-3'. La reconstitution de la séquence codant pour la Met-SAH complète dans pUC8 donne le plasmide pXL868 obtenu de façon similaire au plasmide pXL869.

20      **Exemple 2 : CONSTRUCTION DE VECTEURS DE CLONAGE POUR LA LEVURE**

E.2.1 Isolement et purification du plasmide pKD1

Le plasmide pKD1 peut être purifié à partir d'une culture en fin de phase exponentielle de croissance de la souche de K. drosophilic UCD 51-130 (collection U.C.D., Université de Californie, Davies, CA 95616) selon le protocole suivant qui dérive de celui décrit par Fleer et coll.<sup>37</sup>. Une culture de 1 litre dans le milieu YPD (extrait de levure 1 %, Bacto-peptone (Difco) 2 %, glucose 2 %) est centrifugée, lavée et resuspendue dans une solution de sorbitol 1,2 M, et les cellules sont converties en sphéroplastes en présence de zymolyase (300 µg/ml), d'EDTA 25

- mM, de phosphate 50 mM et de  $\beta$ -mercapto ethanol (1  $\mu$ g/ml). Après lavage dans une solution de sorbitol 1,2 M, les sphéroplastes correspondant à 250 ml de la culture de départ par tube) sont resuspendus dans 2,5 ml de sorbitol 1,2 M et on ajoute le même volume de tampon Tris-HCl 25 mM, glucose 50 mM, EDTA 10 mM, pH 8,0.
- 5 Les étapes suivantes correspondent au protocole de lyse alcaline déjà décrit<sup>32</sup>, à l'exception de la précipitation du surnageant en acétate de potassium qui est faite à 23°C pendant 15 minutes en ajoutant 14 ml d'isopropanol. Le matériel obtenu est traité à la RNase (50  $\mu$ g/ml) à 37°C pendant 20 minutes puis avec la protéinase K (150  $\mu$ g/ml) dans une solution de NaCl 0,5 M, sarkosyl 0,5 % et d'EDTA 25 mM pendant 1
- 10 heure à 60°C. Après centrifugation en microcentrifugeuse Eppendorf, les surnageants sont précipités à l'éthanol pendant 10 minutes à -70°C, puis les culots sont dissous et l'ADN est purifié par centrifugation isopycnique en gradient de CsCl en présence de bromure d'éthidium.

#### E.2.2. Construction du plasmide pCXJ1.

15 L'intermédiaire de construction pUC-URA3 (figure 8) consiste en un fragment HindIII de 1,1 kb contenant le gène URA3 de la levure S. cerevisiae inséré dans le site unique Nari du plasmide pUC19<sup>33</sup>. Le fragment HindIII dérive du plasmide pG63<sup>34</sup> qui a été digéré par HindIII puis traité par le fragment de Klenow de l'ADN polymérase I d'E. coli pour générer des extrémités franches. Le fragment de 1,1 kb contenant le gène URA3 est purifié puis inséré dans le plasmide pUC19

20 lui-même coupé par Nari et traité par le fragment de Klenow de l'ADN polymérase I d'E. coli. Le plasmide pUC-URA3 contient donc : une origine de réPLICATION pour la maintenance du plasmide dans E. coli, le marqueur de résistance à l'ampicilline pour les transformations dans E. coli, le gène lacZ contenant un poly-site de clonage

25 (EcoRI, SacI, KpnI, BamHI, XbaI, SalI, SphI, HindIII comme sites uniques) et le gène URA3 de S. cerevisiae servant de marqueur de sélection dans les mutants uraA de K. lactis.

Le plasmide pCXJ1 (figure 9) contient la séquence complète du plasmide pKD1 insérée dans le site unique AatII de pUC-URA3. Cette construction

a été obtenue en linéarisant pKD1 au site EcoRI puis en traitant ce plasmide par le fragment de Klenow de l'ADN polymérase I d'*E. coli*. Ce fragment d'ADN est ligaturé avec pUC-URA3 précédemment coupé par AatII et traité à la T4 DNA polymérase. La ligature de ces deux fragments permet la reconstitution de sites de restriction EcoRI. L'insertion de l'ADN de pUC-URA3 au site EcoRI du plasmide pKD1 n'inactive aucun gène nécessaire pour maintenir la stabilité et le nombre de copies du plasmide (le site EcoRI étant situé à 205 nucléotides (nt) en amont du codon ATG du gène B<sup>34</sup>). Par conséquent, le plasmide pCXJ1 qui transforme les cellules de K. lactis uraA cir<sup>o</sup> à haute fréquence, est amplifié à environ 70-100 copies par cellule et est maintenu de façon stable même en l'absence de pression de sélection. Grâce à l'origine de réplication apportée par pUC-URA3, le plasmide pCXJ1 peut également se répliquer dans *E. coli* ce qui facilite les étapes de construction et de purification de plasmides. Les sites uniques de pCXJ1 sont les sites HindIII et SalI du polySite de clonage de pUC19.

15      **E.2.3 Construction d'une fusion entre le promoteur pk1 de K. lactis et le gène de la 3'-aminoglycoside phosphotransférase de Tn903**

L'utilisation de pCXJ1 en tant que vecteur pour la transformation de K. lactis et d'autres espèces de Kluyveromyces reste limitée aux souches portant la mutation uraA chromosomique en tant que marqueur d'auxotrophie. Afin de pouvoir transformer des souches sauvages industrielles de Kluyveromyces nous avons choisi le gène de la 3'-aminoglycoside phosphotransférase (aph [3']-I) du transposon bactérien Tn903 comme marqueur dominant de résistance aux antibiotiques que nous avons inséré dans pCXJ1. Le gène de l'aph confère la résistance à la kanamycine chez *E. coli* et rend les souches sauvages de levure résistantes à l'antibiotique G418 (génétidine), lequel est un inhibiteur puissant de la croissance cellulaire<sup>40</sup>. Pour permettre une expression suffisamment forte du gène de l'aph pour avoir un marqueur de sélection dans les transformations de K. lactis, les signaux de transcription bactériens du gène aph sont remplacés par le promoteur pk1 isolé à partir du plasmide «killer» pGkl-1 ou k1 de K. lactis. La construction de la fusion pk1-

aph est effectuée en plusieurs étapes (figures 10 à 12). Tout d'abord un fragment ScaI-PstI de 1,5 kb du plasmide k1 est sous-cloné entre les sites uniques ScaI et PstI de pBR322, le plasmide recombinant sensible à l'ampicilline et résistant à la tétracycline est appelé pk1-PS1535-6 (figure 10). Le fragment sous-cloné de 1,5 kb provient d'une extrémité du plasmide «killer» linéaire et contient la moitié 5' de la première phase de lecture ouverte (ORF1) portée par le plasmide k1 et environ 220 paires de bases en amont<sup>41</sup>. Comme le site ScaI est situé seulement 22 paires de bases de l'extrémité gauche de k1 (figure 10), le fragment ScaI-PstI purifié à partir de gel d'agarose contient vraisemblablement la région promoteur entière de ORF1. La digestion de pk1-PS1535-6 avec DdeI donne un fragment de 266 pb contenant 17 pb provenant de pBR322 à une de ses extrémités (proche de ScaI) et les 11 premiers codons de ORF1 à l'autre extrémité. Après un traitement par le fragment de Klenow de l'ADN polymérase I d'*E. coli*, le fragment purifié est inséré dans le site XbaI unique du plasmide pUC-kan1 (figure 11). Ce dernier plasmide a été obtenu en insérant un fragment EcoRI de 1,25 kb contenant le gène aph provenant de Tn903 (Kanamycin Resistance Gene Block™, Pharmacia) dans le site unique EcoRI de pUC19. Le plasmide pUC-kan202 est obtenu en effectuant une digestion de pUC-kan1 par XbaI suivie d'une digestion brève par la nucléase S1 rendant les extrémités franches puis d'une ligature avec le fragment DdeI de pk1-PS1535-6 rendu franc avec le fragment de Klenow de l'ADN polymérase I d'*E. coli* (figure 11). Cette construction permet d'obtenir une fusion en phase entre les 11 premiers acides aminés du gène ORF1 du plasmide linéaire k1 et de l'extrémité tronquée en 5' du gène aph de Tn903. En fait, la jonction entre ORF1 et aph restore le douzième codon de aph (AGG) de sorte que ce sont les onze premiers acides aminés de aph qui sont remplacés par les onze premiers acides aminés de ORF1 de k1. La séquence complète du promoteur pk1 utilisé pour la fusion pk1-aph ensemble avec le début du gène structural codant pour la aph est indiquée sur la figure 12.

#### E.2.4 Construction du plasmide pKan707

Le plasmide pKan707 (figure 13) est un dérivé du plasmide pCXJ1

- décrit ci-dessus. Il est construit en coupant pCXJ1 par HindIII puis en traitant par le fragment de Klenow de l'ADN polymérase I d'*E. coli*. Le plasmide linéarisé à extrémités franches est ensuite ligaturé avec un fragment ScaI-HincII de 1,2 kb provenant du plasmide pUC-Kan202 portant la fusion pk1'-aph (figure 13). La 5 digestion de pUC-Kan202 par ScaI et HincII donne une cassette de résistance à la kanamycine à extrémités franches ne contenant plus le promoteur bactérien d'origine. Le plasmide obtenu, pKan707, confère la résistance à de très hauts niveaux de G418 (> 2,5 g/l) dans les souches de *K. lactis*. Comme dans le cas de pCXJ1, le plasmide pKan707 peut être transformé dans les souches de *K. lactis cir°* à haute fré- 10 quence, amplifié à 70-100 copies par cellule et maintenu de façon stable en l'absence de pression de sélection (figure 14). Sa grande stabilité conjuguée avec la présence d'un marqueur dominant efficace permettant la transformation de souches industrielles de *Kluyveromyces* font de pKan707 un vecteur de clonage très performant pour les levures du genre *Kluyveromyces*.
- 15      **Exemple 3 : CONSTRUCTION DE VECTEURS D'EXPRESSION CONTENANT DES CASSETTES D'EXPRESSION ET/OU DE SECRETION DE L' ALBUMINE DANS LA LEVURE**
- E.3.1. Construction de cassettes d'expression de la Met-SAH et de la prépro-SAH sous contrôle du promoteur PGK de *S. cerevisiae*.
- 20      Le plasmide pYG12 (figure 15) contient un fragment de restriction SalI-BamHI d'une taille de 1,8 kb constitué des régions promotrices et terminatrices du gène PGK de *S. cerevisiae*. Ce fragment provient d'un fragment génomique HindIII déléte d'un fragment de 1,2 kb correspondant au gène structural et comprenant une région comprise entre l'ATG d'initiation de la traduction et le site BglII 25 localisé 30 codons en amont du codon TAA spécifiant la fin de traduction<sup>42</sup>. Les sites HindIII encadrant le fragment de 1,8 kb ainsi obtenus sont ensuite détruits et remplacés par un site SalI et un site BamHI, respectivement en amont de la région promotrice et en aval du terminateur PGK de transcription. Un site HindIII est introduit à la jonction entre régions promotrices et terminatrices ; le site étant

unique, il permet d'introduire facilement des gènes hétérologues. La séquence nucléotidique de cette région est indiquée sur la figure 15. Le fragment Hind III de 1,8 kb provenant du plasmide pXL868 (voir E.1.3) et codant pour le gène de la Met-SAH est ensuite introduit dans le site HindIII du plasmide pYG12 pour donner le 5 plasmide recombinant pYG10. De façon analogue, le plasmide pYG11 (figure 15) est généré par insertion dans le plasmide pYG12 d'un fragment HindIII provenant du plasmide pXL869 (voir E.1.3) et codant pour la prépro-SAH. En conséquence, deux cassettes d'expression SalI-BamHI d'une taille d'environ 3,6 kb sont ainsi construites, comprenant le promoteur du gène PGK de *S. cerevisiae* ( $P_{PGK}$ ), suivi du gène 10 codant soit pour la Met-SAH soit pour la prépro-SAH et enfin la région correspondant au terminateur de transcription du gène PGK ( $T_{PGK}$ ), région incluant le site de polyadénylation du ARNm.

### E.3.2 Construction des plasmides d'expression pYG19 et pYG23.

Dans le but d'introduire les cassettes d'expressions détaillées ci-dessus, dans le vecteur pKan707 réplicatif chez *Kluyveromyces*, les fragments SalI-BamHI de 3,6 kb et provenant des plasmides pYG10 et pYG11 sont dans un premier temps sous-clonés dans les sites correspondants du plasmide pIC-20<sup>4</sup>, pour donner ainsi les constructions plasmidiques pYG18 (prépro-SAH) (figure 16) et pYG22 (Met-SAH). Ces constructions intermédiaires permettent de disposer des cassettes 20 d'expression  $P_{PGK}$ /Met-SAH/ $T_{PGK}$  et  $P_{PGK}$ /prépro-SAH/ $T_{PGK}$  sous forme de fragments de restriction SalI-SacI directement insérables dans les sites correspondants de pKan707 (figure 17). La digestion de ce dernier par les enzymes SalI et SacI conduit à la suppression du marqueur URA3 ainsi que d'un fragment de 35 pb situé en amont du gène B de pKD1. Les constructions ainsi obtenues appellées pYG19 25 (prépro-SAH) et pYG23 (Met-SAH) (figure 18) comprennent les cassettes d'expression  $P_{PGK}$ /HSA/ $T_{PGK}$ , la séquence quasiment complète du plasmide pKD1 de *K. drosophilaeum*, les séquences permettant la replication autonome chez *E. coli*, ainsi que les gènes codant pour la  $\beta$ -lactamase et la 3'-aminoglycoside phosphotransférase.

(sous contrôle du promoteur *pk1*), permettant respectivement de sélectionner *E. coli* en présence d'ampicilline et *K. lactis* en présence de G418.

**Exemple 4 : TRANSFORMATION DE LA SOUCHE MW98-8C DE *K. lactis* PAR LES PLASMIDES EXPRIMANT LA Met-SAH ET LA prépro-SAH**

5   **E.4.1 Protocole de transformation**

Les plasmides pYG19 et pYG23 sont utilisés pour transformer la souche MW98-8C (*alpha uraA argA lysA K<sup>+</sup> pKD1<sup>o</sup>*) de *K. lactis*<sup>44</sup>. Un échantillon de cette souche a été déposé au Centraalbureau voor Schimmelkulturen (CBS) à Baarn aux Pays Bas selon les dispositions du traité de Budapest où il a été enregistré le / / sous le numéro CBS 88. La transformation s'effectue soit par la technique de formation de sphéroplastes initialement décrite par Hinnen et coll.<sup>44</sup> et adaptée en conséquence, soit en traitant les cellules entières par l'acétate de lithium, ce qui favorise l'incorporation d'ADN en présence de polyéthylène glycol (PEG)<sup>45</sup>. Les modifications concernant la formation de sphéroplastes sont déjà décrites<sup>46</sup>. Lorsque la méthode faisant intervenir l'acétate de lithium est utilisée, la croissance des cellules se fait à 28°C dans 50 ml de milieu YPD, avec agitation et jusqu'à une densité optique à 600 nm (OD<sub>600</sub>) comprise entre 0,6 et 0,8. Les cellules sont récoltées par centrifugation à faible vitesse, lavées dans une solution stérile de TE (10 mM Tris HCl pH 7,4, 1 mM EDTA), resuspendues dans 3-4 ml d'une solution d'acétate de lithium (0,1 M dans du TE) pour donner une densité cellulaire d'environ 2 x 10<sup>8</sup> c/ml, puis incubées à 30°C pendant 1 heure sous agitation modérée.

Des parties aliquotes de 0,1 ml de la suspension résultante de cellules compétentes sont incubées à 30°C pendant 1 heure en présence d'ADN et à une concentration finale de 35 % de PEG4000. Après un choc thermique de 5 minutes à 42°C, les cellules sont lavées 2 fois par de l'eau stérile, resuspendues dans 0,2 ml d'eau stérile et transférées dans des tubes de 10 ml. Une solution d'YPD contenant 0,7 % d'agar et maintenue en surfusion à 46°C (5ml) est ensuite ajoutée et immédiatement coulée sur boîtes YPD. Après modification, une surcouche additionnelle de

5 ml de "top agar" est ajoutée et les boîtes sont incubées à 28°C. Après 18 à 24 heures d'incubation, 0,16 ml d'une solution de G418 (Geneticin 50 mg/ml, GIBCO, Grand Island, N.Y.) est étalée sur les boîtes et des transformants sont comptés après 4-5 jours d'incubation additionnelle à 28°C.

5 L'étalement direct des cellules sur boîtes YPD + G418 (c'est-à-dire sans utiliser une surcouche additionnelle de top-agar) résulte en l'apparition de clones de taille plus importante pour les cellules transformées de *K. lactis*. Cependant, une concentration plus basse de G 418 (50 µg/ml) est alors nécessaire pour observer des colonies, ce qui entraîne la sélection de mutants partiellement résistants au G418 et provenant de cellules non transformées. De plus, l'efficacité de transformation est au minimum 10 fois plus faible que l'efficacité observée après expression phénotypique pendant 18 à 24 heures.

10

#### E.4.2 Stabilité mitotique des plasmides d'expression de la SAH

15 La stabilité mitotique des plasmides pYG19 et pYG23 est mesurée au cours du temps après croissance en milieu non sélectif ; elle est déterminée comme étant le rapport entre les pourcentages finaux et initiaux de cellules croissant sur boîtes YPD contenant 250 µg/ml de G418. Comme l'indique la figure 19, les deux plasmides sont remarquablement stables malgré l'expression à haut niveau d'un gène hétérologue : 40 à 45 % des cellules ont maintenu ces plasmides après 40 générations cellulaires en milieu non sélectif.

20

#### E.4.3 Niveaux d'expression et d'excrétion de l'albumine

Les niveaux d'expression et d'excrétion des cellules de *K. lactis* MW98-8C contenant les plasmides pYG19 (prépro-SAH) et pYG23 (Met-SAH) sont déterminés en fonction du temps après croissance à 28°C en milieu YPD et sous agitation constante. Les surnageants de culture sont obtenus par deux centrifugations consécutives (5 minutes à 4000 puis 10 minutes à 12000 tours/minute) pour éliminer toute contamination éventuelle par des cellules. Un échantillon du second

25

surnageant (0,5 ml) est chauffé à 95°C pendant 15 minutes puis mélangé à un volume égal de tampon d'application contenant 0,125 M Tris HCl, 20 % glycerol, 10 % 2-mercaptopéthanol ( $\beta$ -ME), 4,6 % sodium dodecyl sulfate (SDS) et 0,4 % bleu de bromophénol. Quand une concentration des protéines présentes dans le surnageant  
5 est souhaitée, 0,4 ml d'une solution à 100% p/v d'acide trichloracétique (TCA) est ajoutée à 8 ml de surnageant et les protéines sont précipitées sur glace pendant 1 heure. Le matériel précipité est récupéré par centrifugation (20 minutes, 15 000 tours/minute) puis résolubilisé par chauffage à 95°C pendant 15 minutes dans 0,5 ml  
10 de tampon d'application contenant 63 mM Tris HCl, 10 % glycerol, 5 %  $\beta$ -ME, 2,3 % SDS et 0,2 % bleu de bromophénol.

L'expression intracellulaire d'albumine est détectée à partir d'extraits cellulaires préparés comme suit : 0,25 ml équivalent d'un culot cellulaire (lavé une fois en tampon salin) est resuspendu dans le même volume de tampon de lyse maintenu sur glace et contenant 67 mM de phosphate à pH 7, 5 mM  $\beta$ -ME, 1 mM phénylméthylsulfonyl fluoride (PMSF), 2  $\mu$ M de leupeptine, et 2  $\mu$ M de pepstatine A (Sigma). Après addition de 0,5 ml de billes de zirconium (diamètre 0,45 mm), stockés en tampon phosphate (67 mM, pH 7) à 4°C, les cellules sont cassées par 7 périodes de 30 secondes avec un broyeur de cellules (Biospec Mini Bead-Beater, Biospec Products) entrecoupées par des périodes de refroidissement de 2 minutes  
15 sur glace. L'efficacité de broyage des cellules est supérieure à 90 % si on en juge par comptage des cellules sous microscope à contraste de phase. La fraction liquide dépourvue de billes de zirconium est transférée dans un tube Eppendorf et les billes sont lavées 3 fois avec 0,2 ml de tampon de lyse, puis rassemblées dans le même tube Eppendorf. Par centrifugation du tube pendant 15 minutes à 4°C à 12 000 g, on définit  
20 ainsi une fraction des protéines solubles (surnageant) et une fraction de protéines insolubles (culot). Ces deux fractions sont individualisées et reprises dans le même volume final d'une solution de tampon d'application après dilution. Ces échantillons sont chauffés pendant 15 minutes à 95°C, et appliqués sur un gel de polyacrylamide-  
25

SDS à 8,5 % précédé d'un gel de concentration à 5 %<sup>47</sup>, puis soumis à une intensité de 25 mA jusqu'à ce que le bleu de bromophénol atteigne le bas du gel.

La figure 20 montre le résultat d'une expérience typique qui permet d'évaluer l'expression et l'excrétion d'albumine obtenue avec la souche de *K. lactis* MW 98-8C transformée avec les plasmides pYG19 (prépro-SAH), pYG23 (Met-SAH), et pYG25 (vecteur dépourvu de cassette d'expression), après croissance dans 50 ml de milieu YPD sans G418 à 28°C pendant 68 heures. Chaque échantillon correspond à 100 µl de culture originale et permet de comparer les fractions solubles, les fractions insolubles, et le surnageant de culture, après migration en gel de polyacrylamide à 8,5 % et coloration au bleu de coomassie. Une bande protéique qui co-migre avec de l'albumine humaine commerciale (Sigma) servant de référence de poids moléculaire est détectable dans les échantillons provenant des cellules transformées avec les plasmides pYG19 ou pYG23 mais pas avec les cellules contenant le vecteur pYG25 qui ne contient pas le gène de la SAH. On remarquera que, alors que la quasi-totalité de l'albumine exprimée à partir de pYG19 (prépro-SAH) est exportée dans le surnageant de culture, toute l'albumine produite à partir du plasmide pYG23 (Met-SAH) est présente dans la fraction des protéines cytoplasmiques non solubles. De plus, les résultats de la figure 20 indiquent que l'albumine excretée par les cellules contenant le plasmide pYG19 est la seule protéine extracellulaire présente en quantités significatives. Dans le surnageant d'une culture de MW98-8C transformée avec le plasmide pYG19, la concentration d'albumine peut atteindre 12 mg SAH/g biomasse.

#### E.4.4 Détection immunologique de l'albumine produite par *K. lactis*

L'identité des protéines qui co-migrent avec l'albumine extraite du plasma humain peut être testé par immuno-détection. A ce but, un gel de polyacrylamide est transféré sur filtre de nitrocellulose (Schleicher et Schuell, 0,45 µm) en utilisant un appareil de transfert semi-sec (Biométra) dans un tampon de transfert contenant 25 mM Tris base, 150 mM glycine, et 10 % méthanol. Le temps de transfert

est de 30 minutes en utilisant un courant d'environ 0,85 mA/cm<sup>2</sup> de surface de gel. Après transfert, le filtre est incubé trois fois 5 minutes avec 50 ml de tampon A (5% poudre de lait écrémé, 0,2 % Tween 20 en tampon PBS [137 mM NaCl, 2,7 mM KCl, 4,3 mM Na<sub>2</sub>HPO<sub>4</sub>]), suivi d'une incubation de 30 minutes dans 40 ml de tampon A

5 contenant des anticorps polyclonaux de lapin dirigés contre la SAH et dilués au 1/1000. Après 3 rinçages du filtre par du tampon A, un anticorps secondaire biotinylé reconnaissant les anticorps de lapin (Vectastain ABC - Immuno Peroxidase Kit, Vector Laboratories), est ajouté à raison d'une goutte pour 50 ml de tampon A.

Après incubation du filtre pendant 30 minutes, le filtre est rincé 3 fois par du tampon

10 B (0,2 % Tween 20 dans PBS) puis incubé en présence d'un complexe avidine DH/péroxidase H biotinylée. Le complexe avidine/péroxidase biotinylée est préparé extemporanément en diluant 1 goutte de chacun des réactifs A et B du kit dans 5 ml de tampon B, après incubation à 23°C pendant 30 minutes. Le filtre est incubé dans une dilution au 1/10 du complexe dans du tampon B (50 ml au total) pendant 30

15 minutes. Le filtre est rincé de nouveau par du tampon B pendant 3 périodes de 5 minutes et mis à développer pendant 2-3 minutes dans une solution de 10 ml à 0,02% H<sub>2</sub>O<sub>2</sub> et 10 ml d'une solution contenant 1 mg/ml de diaminobenzidine et 0,4 mg/NiCl<sub>2</sub> en 0,1 M Tris HCl pH 7,4. Le résultat de cette expérience (figure 21) démontre que *K. lactis* MW98-8C exprime (pYG23) et excrète (pYG19) une protéine reconnue par des anticorps polyclonaux anti-SAH. Ce matériel antigénique est spécifiquement associé à la présence d'une cassette d'expression de l'albumine ; il n'est pas détecté dans des extraits cellulaires ou des surnageants de culture des levures transformées avec le vecteur sans cassette d'expression.

20

#### E.4.5 Cinétique d'excrétion de la SAH

25 Comme le montre la figure 22A (chaque échantillon correspondant à l'équivalent de 160 µl de surnageant de culture), la sécrétion de l'albumine humaine se fait avec une cinétique relativement lente en erlenmeyer. Les niveaux maximums

d'excrétion semblent avoir lieu après 70 à 100 heures d'incubation d'une culture inoculée initialement à  $10^5$  cellules/ml (Fig 22B, échantillons déposés correspondant à l'équivalent de 25 µl de surnageant de culture). Aucune augmentation, ni diminution significative n'est détectable entre 100 et 240 heures de culture, démontrant une remarquable stabilité de la SAH dans ces conditions de culture et de température. Puisqu'il n'y a pas d'accumulation de matériel protéolysé durant cette période, il peut être conclu qu'aucune protéase extracellulaire n'est présente qui pourrait dégrader l'albumine excrétée.

Une représentation graphique de la cinétique d'excrétion (A), de même que la courbe de croissance de la souche MW98-8C transformée avec le plasmide pYG19 (B) est montrée en figure 23 ; ces résultats montrent que l'excrétion de l'albumine humaine continue après que les cellules ont atteint la phase stationnaire de croissance. Cette observation est en accord avec des observations antérieures concernant d'autres systèmes d'expression chez la levure \*\*.

15   **Exemple 5 : PURIFICATION DE LA SAH SECRETEE DANS LE MILIEU DE CULTURE**

Dans une expérience typique de purification, la souche MW98-8C transformée avec le plasmide pYG19 est cultivée en milieu YPD pendant 72 heures dans les conditions standards déjà décrites. Un surnageant de culture (0,5 litre) est débarrassé de toute contamination cellulaire par centrifugation et incubé à 4°C pendant 15 minutes en présence de 60 % d'éthanol (V/V). Le précipité est récupéré par centrifugation, redissous dans 10 ml d'une solution 50 mM Tris HCl pH 8,0, puis chargé sur une colonne de bleu Trisacryl. L'albumine humaine est éluee de cette colonne par une solution de NaCl 3M. Après dialyse des fractions contenant la SAH contre 50 µM Tris, celles-ci sont purifiées sur colonne MONO Q (Pharmacia), et élues à une concentration de NaCl 0,25 M. Dans une dernière étape une chromatographie Superose 12 (Pharmacia) permet d'obtenir de la SAH pure à plus de 99% (d'après les résultats d'électrophorèse PAGE-SDS et coloration du gel aux sels

d'argent).

#### Exemple 6 : CARACTERISATION DE L'ALBUMINE SECRETEE ET PURIFIEE

L'absence d'un test permettant de mesurer *in vitro* les propriétés biologiques de l'albumine ne simplifie pas la procédure permettant de juger de la qualité 5 de l'albumine recombinante. Pour cette raison, après purification, l'albumine excrétée par *K. lactis* est donc caractérisée par plusieurs tests physico-chimiques et immunologiques. Ces tests démontrent que l'albumine recombinante est sécrétée par *K. lactis* sous forme mature et dans sa configuration native : elle est indistinguorable de la sérum albumine humaine extraite du plasma pour tous les critères.

##### 10 E.6.1 PAGE SDS : Coloration au bleu de coomassie et aux sels d'argent

L'électrophorèse est effectuée en gel de polyacrylamide (7,5 %) dénaturant (figure 24 C) comme décrit précédemment, ou grâce à l'utilisation d'un "Phast gel" (Pharmacia, Figure 24 A). Différentes quantités d'albumine recombinante ont été comparées à une préparation commerciale de l'albumine standard 15 (Sigma) extraite de plasma humain. La coloration du gel au bleu de coomassie (Figure 24 C), ou aux sels d'argent (Figure 24 A), montre l'homogénéité totale des échantillons d'albumine de levure.

##### E.6.2 Isoélectrofocalisation

Des focalisations isoélectriques sont effectuées entre pH 4,5 et 6,0 20 (Phast gel, Pharmacia, figure 24 B), et pH 4,0 et 7,0 (Immobiline, LKB). Le point isoélectrique de la SAH recombinante est identique à celui de la SAH humaine référence ( $pI = 4,8$ ).

##### E.6.3 PAGE en conditions non dénaturantes : révélation immunologique

L'électrophorèse de la SAH recombinante dans un gel de polyacrylate 25 (10 %) non dénaturant, suivi d'un transfert sur filtre de nitrocellulose et

révélation immunologique dans les conditions décrites en section E.4.4 révèlent une co-migration avec l'albumine standard extraite du plasma humain. Ceci suggère très fortement que la maturation correcte de la prépro-SAH recombinante a lieu lors du processus de sécrétion inhérent aux cellules de K. lactis. Aucune contamination avec de la SAH non maturée n'est détectable ce qui implique que le clivage de la séquence signal et de la région pro de la SAH s'effectue avec efficacité dans cette levure.

#### E.6.4 Chromatographie d'exclusion moléculaire

Une chromatographie d'exclusion moléculaire est effectuée avec une Superose 12 (Pharmacia). Le débit du tampon d'élution (50 mM Tris HCl pH 8,0) est maintenu à 1 ml/minute. Les concentrations en albumine recombinant et en albumine référence issue du plasma humain sont de 0,4 et 1 mg/ml respectivement. Après dépôt d'un volume de 100 µl, l'élution de l'albumine est détectée par mesure de l'absorption à 280 nm. Dans les deux cas le volume d'élution est identique (11,5 ml).

#### E.6.5 Chromatographie d'échange d'anion

L'albumine recombinante est éluee d'une colonne Mono Q HR 5/5 (Pharmacia) à une concentration de 0,31 M NaCl, comme l'albumine référence issue du plasma humain. La figure 25 montre les profils chromatographiques obtenus pour des injections de 100 µl de SAH recombinante (0,4 mg/ml) et de SAH plasmatique (0,5 mg/ml) utilisant une colonne équilibrée en Tris-HCl 50 mM pH 8,0 et à un débit constant de 1 ml/minute.

#### E.6.6 Chromatographie en phase inverse

Le comportement de l'albumine de levure en phase inverse est analysé sur une colonne Nucleosil (C4) avec les tampons A (eau à 0,1 % TFA) et B (acetonitrile à 0,1 % TFA). La figure 26 représente l'élution, par un gradient de 20 à 80 % de B dans A, de l'albumine de levure qui possède le même temps de rétention

que l'albumine plasmatique.

#### E.6.7 Composition en acides aminés

La composition en acides aminés de la SAH recombinante est déterminée par chromatographie en phase inverse après hydrolyse acide (HCl) et 5 dérivatisation au PITC. Les résultats obtenus par cette méthode indiquent clairement une composition identique à celle de l'albumine référence issue du plasma humain.

#### E.6.8 Séquence N-terminale

L'utilisation d'un appareil automatisé permettant la dégradation 10 d'Edman (Applied Biosystems) indique que la séquence N-terminale de la SAH secrétée par K. lactis est Asp-Ala-His-Lys-Ser-Glu-Val-Ala-His-Arg-Phe-Lys-Asp-Leu-Gly. La détermination de cette séquence confirme donc les résultats provenant des expériences d'électrophorèse en gels natifs de polyacrylamide, et démontre la maturation correcte et complète de l'albumine secrétée par la levure K. lactis.

#### 15 E.6.9 Fluorescence du tryptophane

Avec une excitation à 295 nm, l'émission de fluorescence de l'unique tryptophane possède le même maximum pour la SAH recombinante et l'albumine plasmatique (337 nm) indiquant un environnement hydrophobe identique au niveau de cet acide aminé (figure 27).

#### 20 E.6.10 Stabilité thermique

L'albumine référence (Sigma) et l'albumine secrétée par la levure ont une cinétique identique de dégradation à 75°C (figure 28) ce qui suggère que les liaisons de stabilisation de la structure protéique (interactions hydrophobes et ponts disulfures) sont identiques dans les deux cas.

### E.6.11 Affinités vis-à-vis de différents anticorps monoclonaux

L'antigénicité de l'albumine recombinante a été étudiée avec différents anticorps monoclonaux dirigés contre l'albumine humaine. La figure 29 montre leur spécificité : les anticorps HA10, HA11 et HA13 correspondent à des épitopes dont la conservation nécessite l'intégrité de l'ensemble de la molécule. Les anticorps HA21, HA6, HA4, HA3, HA8 et HA1 correspondent à des épitopes localisés le long de la chaîne peptidique de l'extrémité N à l'extrémité C terminale<sup>48, 50</sup>.

Le test utilisé est un test d'inhibition d'ELISA. Les différents anticorps sont adsorbés sur des plaques de polystyrène et une courbe de fixation d'albumine marquée avec la phosphatase alcaline est établie pour chaque anticorps. Les courbes de saturation de chaque anticorps ont une allure sigmoïde et la quantité d'albumine marquée correspondant à 50 % de fixation a été choisie pour étudier l'inhibition de chacun des anticorps.

L'inhibition a été réalisée avec l'échantillon d'albumine recombinante et comparée à celle obtenue avec un échantillon d'albumine plasmatique (Sigma). La figure 30 montre que pour les 9 anticorps testés, l'albumine recombinante est aussi inhibitrice que l'albumine native. Etant donné l'extrême sensibilité du test, les différences observées sont minimales.

### DESCRIPTION DES FIGURES

20 Les représentations des plasmides indiqués dans les figures ne sont pas tracées à l'échelle et seuls les sites de restriction importants pour les constructions sont indiqués.

25 Figure 1 : Construction du plasmide pXL276 contenant la séquence complète codant pour la Met-SAH obtenu à partir de trois clones d'ADNc (voir texte) dérivés d'ARNm poly-A isolés de foie humain.

Figure 2 : Reconstitution de la séquence signal "prépro" de l'albumine à partir de quatre oligonucléotides synthétiques : Construction du plasmide pXL290.

Figure 3 : Construction du plasmide pXL299.

- Figure 4 : Construction du plasmide pXL322.
- Figure 5 : Construction du plasmide pXL855.
- Figure 6 : Construction du plasmide pXL869.
- 5 Figure 7 : Séquence nucléotidique et en acides aminés du fragment HindIII provenant du plasmide pXL869 et contenant le gène de structure de la prépro-SAH. Les flèches noires indiquent la fin des régions "pré" et "pro".
- Figure 8 : Construction du plasmide pUC-URA3.
- Figure 9 : Construction du plasmide pCXJ1.
- Figure 10 : Construction du plasmide pk1-PS1535-6.
- 10 Figure 11 : Construction du plasmide pUC-kan202.
- Figure 12 : Séquence nucléotidique du promoteur pk1 à partir du site de restriction Scal et incluant la jonction entre ORF1 et pk1 et la région 5' du gène bactérien codant pour l'aminoglycoside phosphotransférase provenant de Tn903.
- 15 Figure 13 : Construction du plasmide pKan707.
- Figure 14 : Courbe de stabilité du plasmide pKan707 dans la souche MW98-8C dans des conditions de croissance non-sélectives.
- Figure 15 : Construction du plasmide pYG11.
- Figure 16 : Construction du plasmide pYG18.
- 20 Figure 17 : Construction du plasmide pYG19.
- Figure 18 : Représentation des plasmides pYG19 (prépro-SAH) et pYG23 (Met-SAH). La séquence nucléotidique fournie en dessous de chaque plasmide indique la jonction entre le promoteur du gène PGK et les gènes de structure des différentes formes de SAH.
- 25 Figure 19 : Courbes de stabilité des plasmides pYG19 et pYG23 dans la souche MW98-8C dans des conditions de croissance non-sélectives.
- Figure 20 : Gel de polyacrylamide-SDS à 8,5 % après coloration au bleu de coomassie démontrant l'expression et l'excrétion de la SAH à partir de la souche MW98-8C. Pistes 1 - 4 : albumine extrait du plasma humain (Sigma) déposée à des concentrations croissantes ; 0,2 µg, 0,6 µg, 2 µg et 6 µg par piste. pYG19 (prépro-SAH), pYG23 (Met-SAH) et pYG25 (vecteur) : chaque piste correspond à une quantité de protéine équivalent à 100 µl de la culture originale (approx.  $2 \times 10^8$  cellules/ml). a), fractions solubles ; b), fractions insolubles ; c), surnageant de culture.
- 30
- 35 Figure 21 : Immunorévélation d'un gel de polyacrylamide à 7,5 % transféré sur nitrocellulose comme décrit dans le texte. Piste 1 - 5 : albumine extraite du plasma humain (Sigma) qui a été dissous dans un milieu YPD à des concentrations variées et précipité avec du TCA (5 % concentration finale) comme décrit. Les échantillons de la gamme SAH ont donc été

traités de la même façon que les surnageants de culture. Dans les deux cas, les dépôts de gel correspondent à des volumes équivalents. 1), précipitation à une concentration de 100 mg/l ; 2), 50 mg/l ; 3), 25 mg/l ; 4), 12,5 mg/l ; 5), 0,6 mg/l. pYG19 (prépro-SAH) et pYG23 (Met-SAH) : chaque piste correspond à une quantité de protéines équivalent à 20 µl de la culture originale (approx.  $2 \times 10^8$  cellules/ml). a), surnageant de culture ; b), fractions insolubles ; c), fractions solubles.

**Figure 22 :** Gel de polyacrylamide à 8,5 % après coloration au bleu de coomassie démontrant la cinétique de l'excrétion de la SAH à partir de la souche MW98-8C.

A, piste a-e : albumine extraite du plasma humain (Sigma) qui a été dissous dans du milieu YPD à des concentrations variées et précipitée avec du TCA (5 % concentration final) comme décrit (voir légende de la figure 21). 1), précipitation à une concentration de 6 mg/l ; 2), 12,5 mg/l ; 3), 25 mg/l ; 4), 50 mg/l ; 5), 100 mg/l. pYG19 : excrétion de la SAH à partir du plasmide pYG19 et en fonction de l'âge de la culture, 16 - 61 heures comme indiqué. Chaque échantillon correspond à l'équivalent de 160 µl de surnageant de culture.

B, pistes a -e : albumine extraite du plasma humain (Sigma) déposée à des concentrations croissantes ; 0,4 µg, 0,6 µg, 0,8 µg, 1,0 µg et 1,2 µg par piste. pYG19 : excrétion de la SAH à partir du plasmide pYG19 et en fonction de l'âge de la culture, 72-240 heures comme indiqué. Chaque échantillon correspond à l'équivalent de 25 µl de surnageant de culture.

**Figure 23 :** A, représentation graphique de la cinétique d'excrétion de la SAH à partir de la souche MW98-8C transformée avec le plasmide pYG19 et cultivée en erlenmeyer comme indiqué dans le texte. La concentration de l'albumine détectée dans le milieu de culture (mg/l) est représentée en fonction du temps (heures). Les résultats de quatre expériences indépendantes sont représentés.

B, courbe de croissance de la souche MW98-8C transformée avec le plasmide pYG19.

**Figure 24 :** Techniques électrophorétiques (S = standard, L = albumine de levure).

A - PAGE-SDS (Phast gel, Pharmacia, gradient 8-25 %) coloration aux sels d'argent. S = 1 µg, L : de 0,05 à 0,25 µg

B - Isoélectrofocalisation (Phast gel, pH 4,5-6,0), coloration au bleu de coomassie, dépôts = 1 µg.

C - PAGE SDS (7,5 %), coloration au bleu de coomassie.

S = 0,2 - 1,0 µg, L = 0,25-2,5 µg

**Figure 25 :** Chromatographie d'échange d'ions = Mono Q (Pharmacia). En bas:

- SAH levure 40 µg ; en haut : SAH standard 50 µg
- Figure 26 : Chromatographie en phase inverse = Nucleosil C4. a), SAH standard ; b), SAH levure.
- Figure 27 : Spectrométrie d'émission de fluorescence du tryptophane. En trait plein : SAH standard ; en pointillé : SAH levure.
- Figure 28 : Stabilité thermique à 75°C. En trait plein : SAH standard ; en pointillé : SAH levure.
- Figure 29 : Régions de l'albumine reconnues par les différents anticorps monoclonaux utilisés pour la caractérisation antigénique.
- 10 Figure 30 : Courbes d'inhibition, vis-à-vis des neuf anticorps monoclonaux (voir texte) : pourcentage d'inhibition en fonction de la concentration en albumine (µg/ml) : □, SAH standard ; ■, SAH levure.

#### REFERENCES

- 1 Brown, J.R., in "Albumin Structure, Function and Uses", Rosenoer, V.M. et al. (eds.), Pergamon Press, Oxford, (1977) 27 - 51.
- 15 2 Weitkamp, L.R. et al., Ann. Hum. Genet. 37 (1973) 219 - 226.
- 3 Judah, J.D., and Quinn, P.L., Nature 271 (1987) 384 - 385.
- 4 Bathurst, I.C. et al., Science 235 (1987) 348 - 350.
- 5 Takahashi, N. et al., Proc. Natl. Acad. Sci. USA 84 (1987) 7403 - 7407.
- 20 6 Brennan, S.O. and Carrel, R.W., Nature 274 (1978) 908 - 909.
- 7 Abdo, Y. et al., FEBS Letters 131 (1981) 286 - 288.
- 8 Nagata et al., Rinsho Byori 30 (1982) 791 - 796.
- 9 Douglass, O. et al., Annu. Rev. Biochem. 53 (1984) 665 pp.
- 10 Julius, D.J. et al., Cell 32 (1983) 839 pp.
- 25 11 Wésolowski-Louvel, M. et al., Yeast 4 (1988) 71 - 81.
- 12 Cohn, E.J. et al., J. Am. Chem. Soc. 68 (1946) 459 pp.
- 13 Liautaud, J. et al., 13th Intern. Congress of IABS, Budapest; A: "Purification of proteins. Development of biological standard", Karger (ed.), Bale, 27 (1973) 107 pp.
- 30 14 Klausner, A., Biotechnology 3 (1985) 119 - 125.
- 15 Latta, M. et al., Bio/Technology 5 (1987) 1309 - 1314.
- 16 European patent application EP 198 745, publ. 22.10.1986
- 17 European patent application EP 236 210, publ. 09.09.1987
- 18 Saunders, C.W. et al., J. Bacteriol. 169 (1987) 2917 - 2925.
- 35 19 European patent application EP 0 229 712 A2, publ. 22.07.1987
- 20 Etcheverry, T. et al., Bio/Technology 4 (1986) 726 - 730.
- 21 European patent application EP 0 201 239, publ. 12.11.1986
- 22 Das, S. and Hollenberg, C.P., Current Genetics 6 (1982) 123 - 128.

- 23 International patent application WO 83/04050, publ. 24.11.1983  
24 International patent application WO 83/04051, publ. 24.11.1983  
25 de Louvencourt, L. et al., *J. Bacteriol.* 154 (1982) 737 - 742.  
26 European patent application EP 0 095 986, publ. 07.12.1983  
5 27 European patent application EP 0 241 435 A2, publ. 14.10.1987  
28 Futcher, A.B., *Yeast* 4 (1988) 27 - 40.  
29 Falcone, C. et al., *Plasmid* 15 (1986) 248 - 252.  
30 Chen, X.J. et al., *Nucl. Acids Res.* 14 (1986) 4471 - 4481.  
31 Sherman, F. et al., "Methods in Yeast Genetics", Cold Spring Harbor  
10 Laboratory, Cold Spring Harbor, N.Y., 1986  
32 Maniatis, T. et al., "Molecular Cloning: a Laboratory Manual", Cold  
Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1982  
33 Ausubel, F.M. et al. (eds.), "Current Protocols in Molecular Biology",  
John Wiley & Sons, New York 1987  
15 34 Taylor, J.W. et al., *Nucleic Acids Res.* 13 (1985) 8749 - 8764.  
35 Birnboim, H.C. and Doly, J., *Nucleic Acids Res.* 7 (1979) 1513 - 1523.  
36 Holmes, D.S. and Quigley, M., *Anal. Biochem.* 114 (1981) 193 - 197.  
37 Fleer et al., *Mol. Cell. Biol.* 7 (1987) 1180 - 1192.  
38 Yanisch-Perron, C. et al., *Gene* 33 (1985) 103 - 119.  
20 39 Gerbaud, C. et al., *Curr. Genet.* 3 (1981) 173 - 180.  
40 Jimenez, A. and Davis, J., *Nature* 287 (1980) 869 - 871.  
41 Sor, F. and Fukuhara, H., *Curr. Genet.* 9 (1985) 147 - 155.  
42 Mellor, J. et al., *Gene* 24 (1983) 1 - 14.  
43 Marsh, L. et al., *Gene* 32 (1984) 481 - 485.  
25 44 Hinnen, A. et al., *Proc. Natl. Acad. Sci. USA* 75 (1978) 1929 - 1933.  
45 Ito, H. et al., *J. Bacteriol.* 153 (1983) 163 - 168.  
46 Bianchi, M. et al., *Curr. Genet.* 12 (1987) 185 - 192.  
47 Laemmli, U.K., *Nature* 227 (1970) 680 - 685.  
48 Tschopp, J.F. et al., *Bio/Technology* 5 (1987) 1305 - 1308.  
30 49 Doyen, N. et al., *Mol. Immun.* 22 (1985) 1 - 10.  
50 Lapresle, C., *Anal. Biochem.* (1988) in press.

## REVENDICATIONS

- 1 - Procédé de préparation microbiologique de la sérum albumine humaine (SAH) ou de l'un de ses variants caractérisé en ce que l'on cultive une levure capable d'assurer le maintien stable d'une cassette d'expression comportant au moins un 5 marqueur de sélection des levures transformées, l'ADN du gène de structure de SAH fusionné à des séquences permettant son expression dans la levure et éventuellement l'excrétion de la protéine encodée par ce gène dans le milieu de la culture.
- 2 - Procédé selon revendication 1 caractérisé en ce que la cassette d'expression est intégrée dans le génome de la levure.
- 10 3 - Procédé selon revendication 1 caractérisé en ce que la cassette d'expression fait partie d'un plasmide comportant un système de réPLICATION fonctionnant dans la levure et assurant le maintien stable de ladite cassette dans cette levure.
- 4- Procédé selon l'une des revendications 1, 2 ou 3 caractérisé en ce que les levures sont choisies dans les genres Saccharomyces et Kluyveromyces.
- 15 5- Procédé selon l'une des revendications 1, 2 ou 3 caractérisé en ce que la levure est choisie dans le genre Kluyveromyces.
- 6 - Procédé selon l'une des revendications 1, 2 ou 3 caractérisé en ce que la levure est choisie parmi toutes les variétés de Kluyveromyces marxianus.
- 7 - Procédé selon l'une des revendications 1, 2 ou 3 caractérisé en ce que la 20 levure est Kluyveromyces marxianus var. lactis.
- 8 - Procédé selon l'une des revendications 3 à 7 caractérisé en ce que le système de réPLICATION fonctionnant dans la levure est tout ou partie du plasmide pKD1 isolé de Kluyveromyces marxianus var. drosophilae ou tout ou partie du plasmide 2  $\mu$  isolé de Saccharomyces cerevisiae ou une combinaison d'éléments dérivés du 25 plasmide pKD1 et du plasmide 2  $\mu$ .
- 9 - Procédé selon l'une des revendications 3, 6, 7 ou 8 caractérisé en ce que le système de replication fonctionnant dans la levure est tout ou partie de la séquence du plasmide pKD1.
- 10- Procédé selon l'une des revendications 3, 5, 6, 7, 8 ou 9 caractérisé en ce

qu'une cassette d'expression selon la revendication 1 ou un marqueur de sélection est inséré dans une région de 197 pb définie par les sites SacI et MstII de pKD1.

11 - Procédé selon l'une des revendications 1 à 10 caractérisé en ce que les séquences permettant l'expression du gène de structure sont choisies parmi les 5 promoteurs dérivées des gènes de levures du genre Saccharomyces ou Kluyveromyces.

12 - Procédé selon revendication 11 caractérisé en ce que ces promoteurs sont dérivés des gènes glycolytiques de levures du genre Saccharomyces ou Kluyveromyces.

10 13 - Procédé selon revendication 12 caractérisé en ce que les séquences permettant l'expression du gène de structure sont choisies parmi les gènes codant pour la phosphoglycerate kinase (PGK), la glycéraldéhyde-3-phosphate deshydrogénase, la lactase, l'enolase ou l'alcool deshydrogénase.

15 14 - Procédé selon revendication 13 caractérisé en ce que les séquences permettant l'expression du gène de structure proviennent du gène codant pour la phosphoglycérat kinase.

20 15 - Procédé selon l'une des revendications 1 à 14 caractérisé en ce que les séquences permettant l'excrétion de la protéine hétérologue sont choisies parmi l'extension N-terminale naturelle "prépro" de l'albumine et les séquences obtenues à partir des gènes de levure codant pour la phéromone alpha ou la toxine "killer".

16 - Procédé selon revendication 15 caractérisé en ce que la séquence permettant l'excrétion de la SAH est l'extension terminale naturelle "prépro" de l'albumine.

25 17 - Procédé selon l'une des revendications 1 à 16 caractérisé en ce que les séquences permettant la sélection des levures transformées sont choisies parmi les gènes apportant la résistance aux antibiotiques ou aux ions cuivre.

18 - Procédé selon revendication 17 caractérisé en ce que la séquence de sélection est un gène apportant la résistance à la G418.

30 19 - Procédé selon revendication 18 caractérisé en ce que la séquence de sélection est une fusion entre le promoteur k1 du plasmide linéaire de K. lactis et

l'aminoglycoside phosphotransférase du transposon Tn903.

20 - Le plasmide pYG19 caractérisé en ce qu'il comporte le plasmide pKD1, l'ADN de la prépro-SAH, les séquences d'initiation et d'arrêt de la transcription du gène codant pour la phosphoglycérate kinase et les séquences d'une fusion entre le 5 promoteur pk1 du plasmide linéaire k1 de Kluyveromyces lactis et le gène de la 3'-aminoglycoside phosphotransférase de Tn903.

21 - Le plasmide pYG23 caractérisé en ce qu'il comporte le plasmide pKD1, l'ADN de la Met-SAH, les séquences d'initiation et d'arrêt de la transcription du gène codant pour la phosphoglycerate kinase et les séquences d'une fusion entre le 10 promoteur pk1 du plasmide linéaire k1 de K. marxianus var. lactis et le gène de la 3'-aminoglycoside phosphotransférase de Tn903.

22 - La levure K. marxianus var. lactis MW98-8C.

23 - La levure K. marxianus var. lactis MW98-8C transformée par le plasmide pYG19.

15 24 - La levure Kluyveromyces marxianus var. lactis MW98-8C transformée par le plasmide pYG23.

25 - La sérum albumine humaine lorsqu'elle est obtenue par le procédé selon l'une des revendications 1 à 19.

26 - La sérum albumine humaine lorsqu'elle est obtenue par culture dans des 20 conditions appropriées de la levure selon les revendications 23 et 24.

27 - Application de la sérum albumine humaine selon les revendications 25 ou 26 à titre de médicament.

28 - Le plasmide pKan-707.

29 - Le plasmide CXJ1.

2635115

Pl I/31

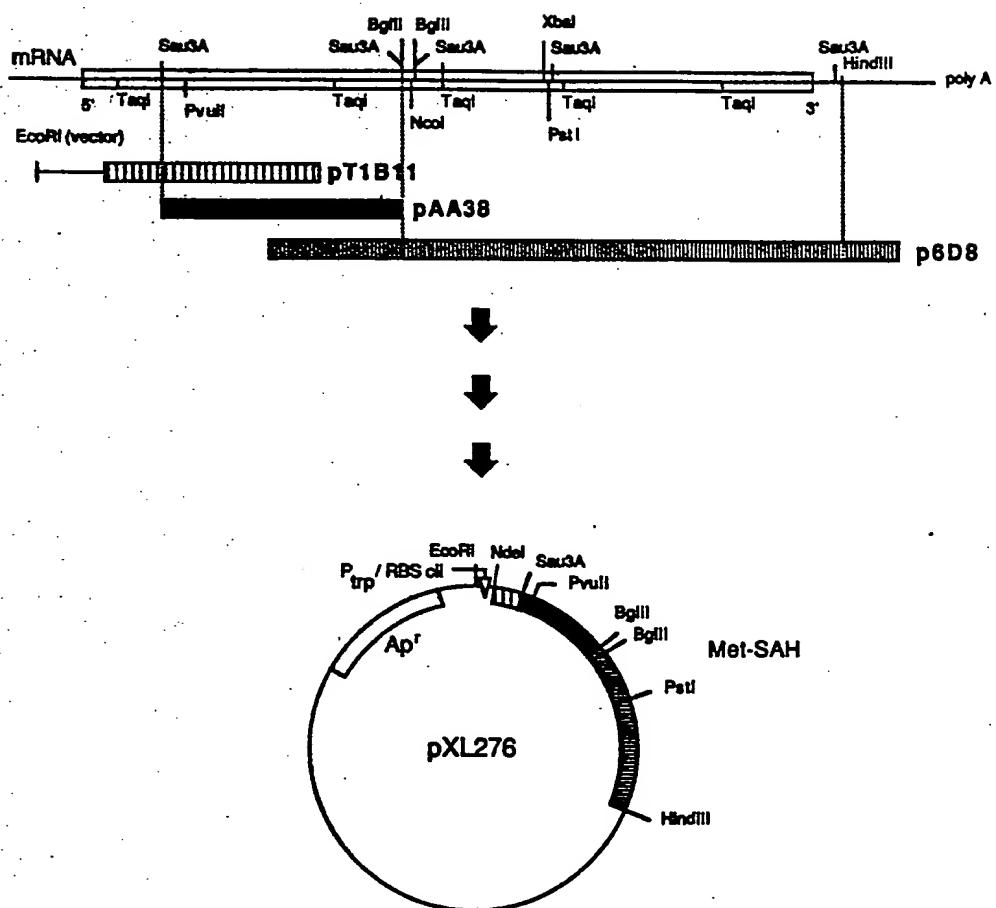


Figure 1

## Pl. II/31

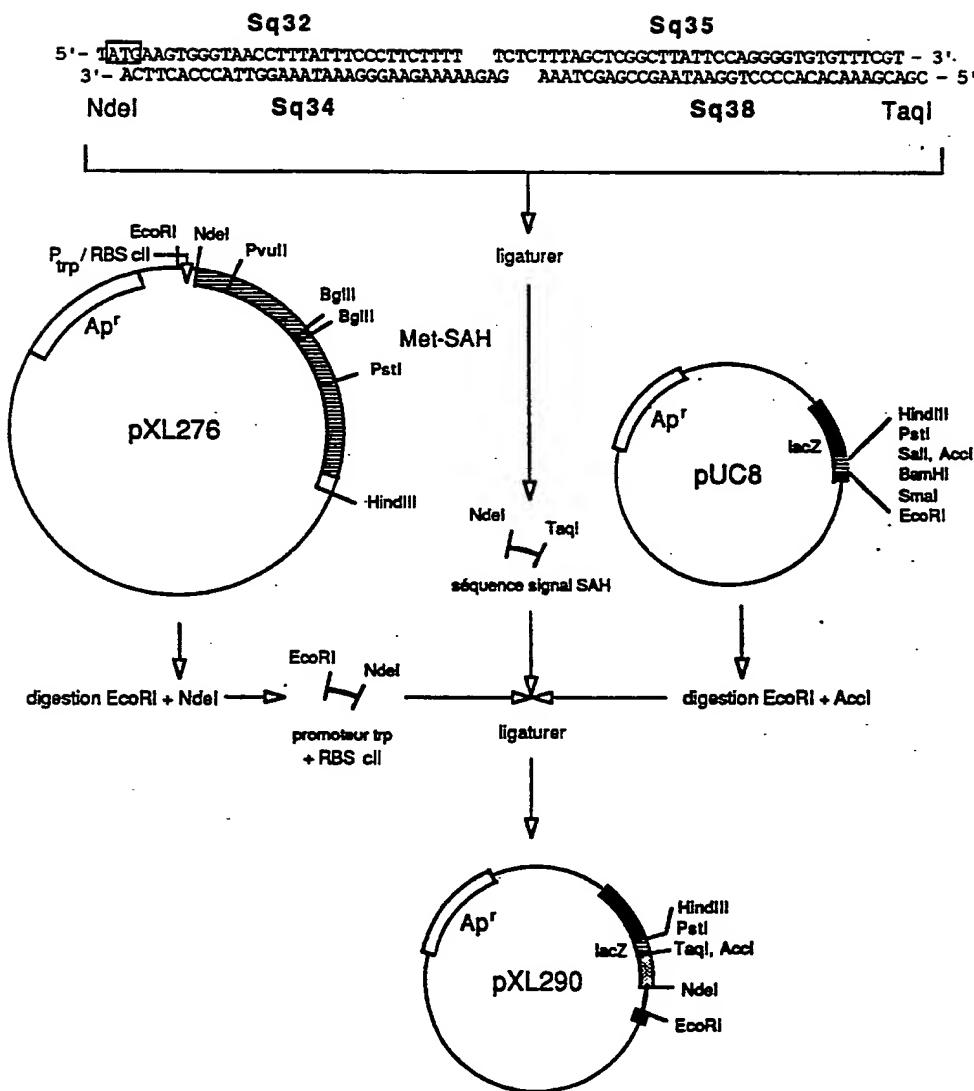


Figure 2

## Pl. III/31

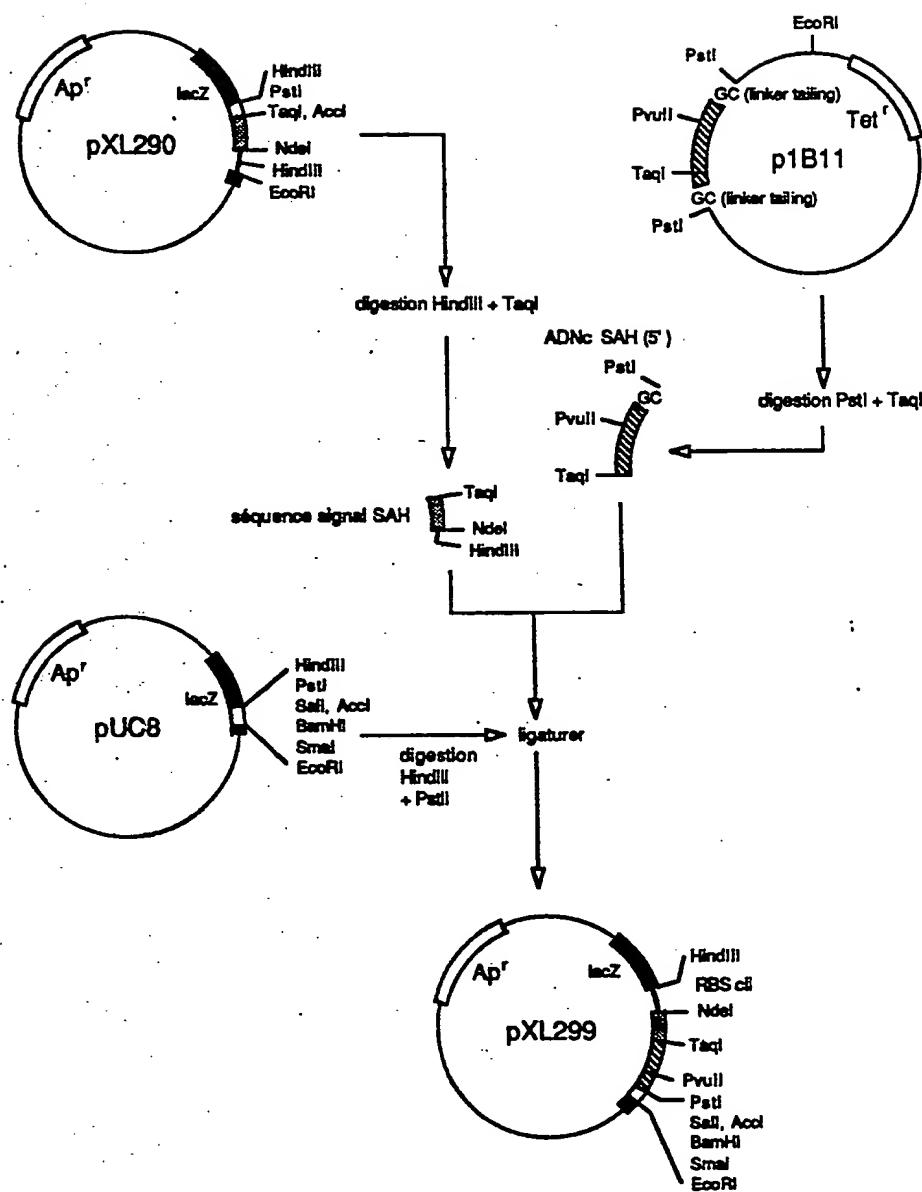


Figure 3

## PL IV/31

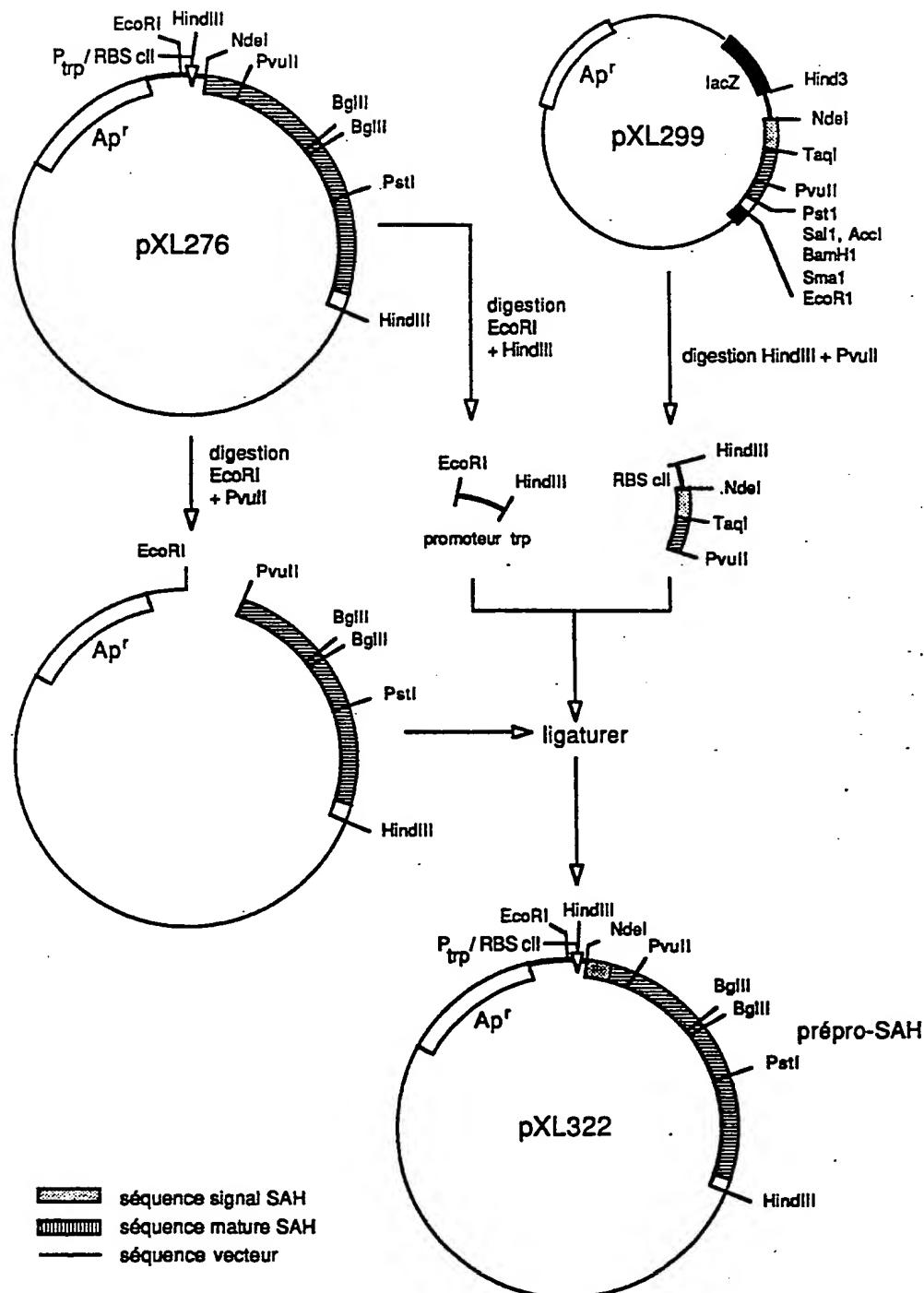


Figure 4

## PL V/31

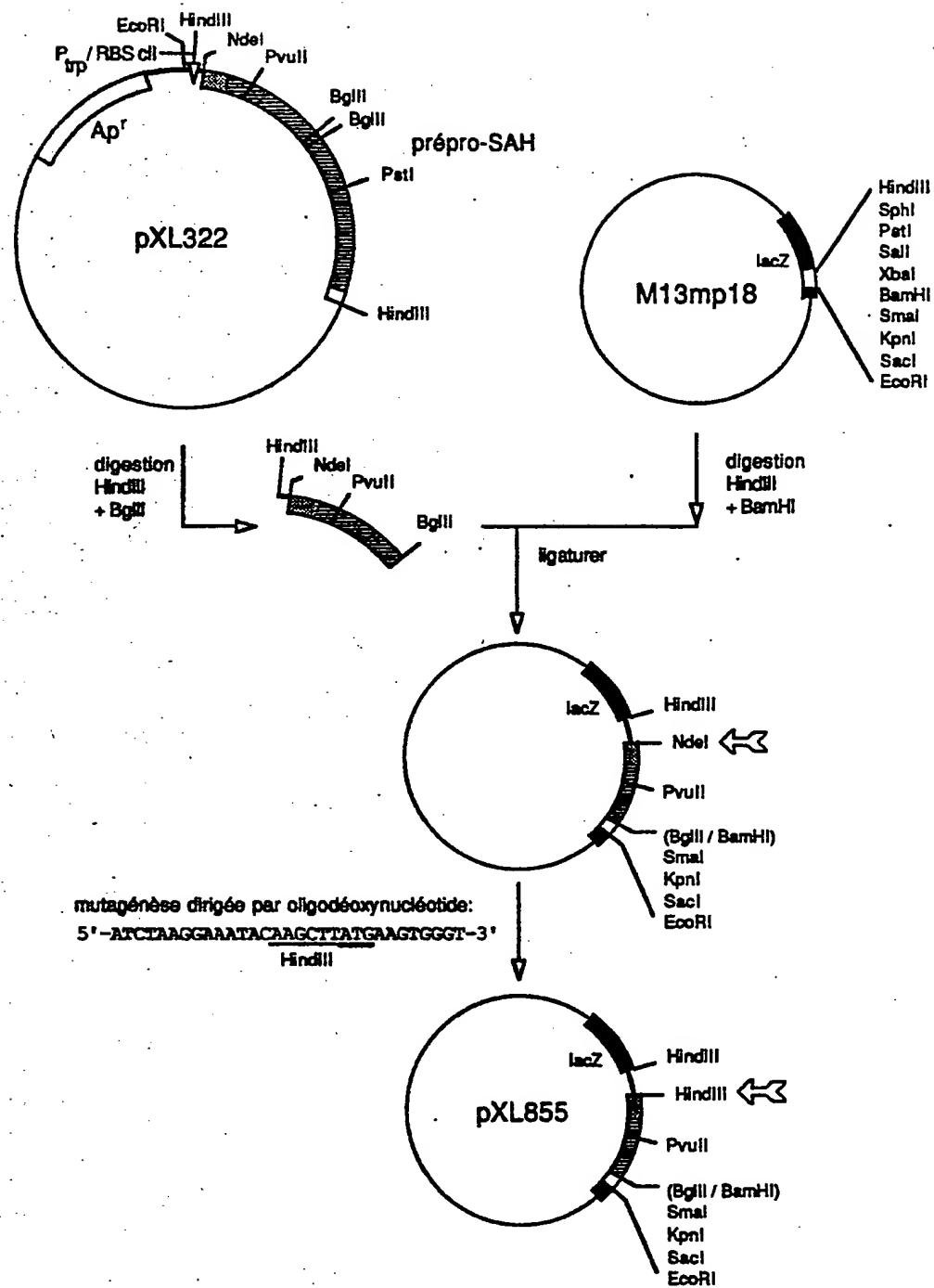


Figure 5

## Pl. VI/31

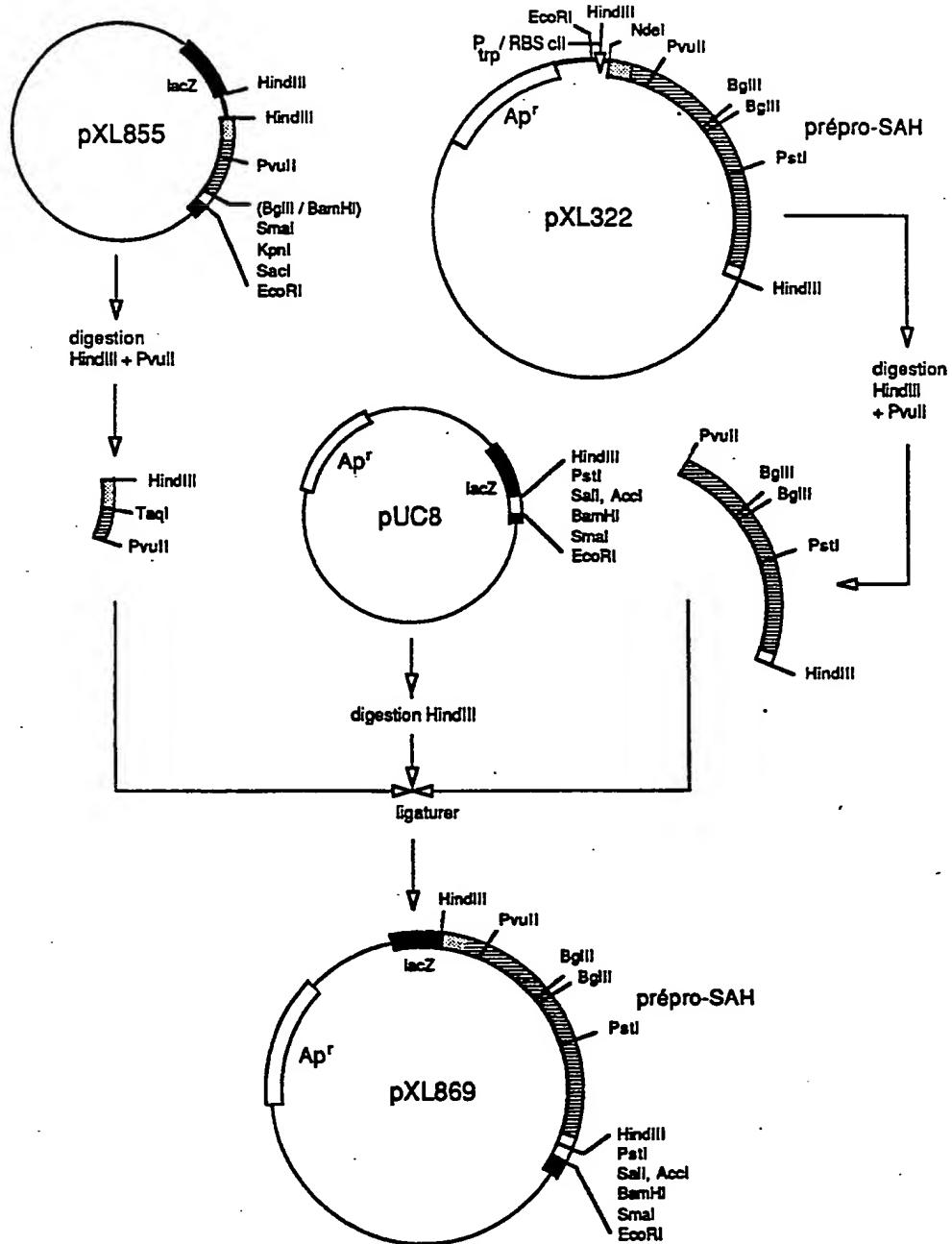


Figure 6

## Pl VII/31

MetLysTrpValThrPheIleSerLeuLeuPheSerSerAlaTyrSerArgGlyValPheArg  
 1 11 21 31 41 51 61 71  
 AAGCTTATGAAGTGGGTAACTTTATTCCTCTTTCTCTGGCTTATTCCAGGGTGTGTTCGT  
  
 ArgAspAlaHisLysSerGluValAlaHisArgPheLysAspLeuGlyGluGluAsnPheLysAlaLeuValLeu  
 CGAGATGCACACAAGAGTGAGGTTGCTCATCGTTAAAGATTTGGGAGAAGAAAATTCAAAGCCTGGTGTGTTG  
 76 86 96 106 116 126 136 146  
  
 IleAlaPheAlaGlnTyrLeuGlnGlnCysProPheGluAspHisValLysLeuValAsnGluValThrGluPhe  
 ATTCCTTGCAGTATCTCAGCTGTCCATTGAGATCATGAAATTAGTAGAATGAAGTAAGTGAATT  
 151 161 171 181 191 201 211 221  
  
 AlaLysThrCysValAlaAspGluSerAlaGluAsnCysAspLysSerLeuHisThrLeuPheGlyAspLysLeu  
 GCAAAAAACATGTTGCTGATGAGTCAGCTGAAAATTGTCACAAATCACTCACCCCTTTGGAGACAAATT  
 226 236 246 256 266 276 286 296  
  
 CysThrValAlaThrLeuArgGluThrTyrGlyGluMetAlaAspCysCysAlaLysGlnGluProGluArgAsn  
 TGACAGTTGCAACTCTCGTGAAACCTATGGTGAATGGCTGACTGCTGTGCAAAACAAGAACCTGAGAGAAAT  
 301 311 321 331 341 351 361 371  
  
 GluCysPheLeuGlnHisLysAspAspAsnProAsnLeuProArgLeuValArgProGluValAspValMetCys  
 GAATGCCCTGCAACACAAAGATGACAATCCAATCTCCCCGATGGTGAAGACCAGAGGTTGATGTGATGTGC  
 376 386 396 406 416 426 436 446  
  
 ThrAlaPheHisAspAsnGluGluThrPheLeuLysLysTyrLeuTyrGluIleAlaArgArgHisProTyrPhe  
 ACTGCTTTCATGACAATGAAGAGACATTGGAAATTGCAAGACATCCCTTACTT  
 451 461 471 481 491 501 511 521  
  
 TyrAlaProGluLeuLeuPhePheAlaLysArgTyrLysAlaAlaPheThrGluCysCysGlnAlaAlaAspLys  
 TATGCCCGGAACCTCTTCTTGCTAAAGGTATAAACCTGCTTTACAGAATGTGCCAAGCTGCTGATAAAA  
 526 536 546 556 566 576 586 596  
  
 AlaAlaCysLeuLeuProLysLeuAspGluLeuArgAspGluGlyLysAlaSerSerAlaLysGlnArgLeuLys  
 GCAGCCTGCCCTGTCACGCTCGATGAACTTCGGGATGAGGGACGCTCGCTGCCAACAGAGACTCAAG  
 601 611 621 631 641 651 661 671  
  
 CysAlaSerLeuGlnLysPheGlyGluArgAlaPheLysAlaTrpAlaValAlaArgLeuSerGlnArgPhePro  
 TGTGCCAGTCTCCAAAATTGGAGAAAGAGCTTCAAAACATGGCAGTAGCTGCCCTGAGCCAGAGATTCCC  
 676 686 696 706 716 726 736 746  
  
 LysAlaGluPheAlaGluValSerLysLeuValThrAspLeuThrLysValHisThrGluCysCysHisGlyAsp  
 AAAGCTGAGTTGCAAGAGTTCAAGTTAGTGACAGATCTTACCAAGTCCACCGAAATGCTGOCATGGAGAT  
 751 761 771 781 791 801 811 821  
  
 LeuLeuGluCysAlaAspAspArgAlaAspLeuAlaLysTyrIleCysGluAsnGlnAspSerIleSerSerLys  
 CTGCTGAAATGCGCTGATGACAGGGCGAACCTTGCAAGTATATCTGAAATCAAGATTGATCTCCAGTAAA  
 826 836 846 856 866 876 886 896  
  
 LeuLysGluCysCysGluLysProLeuLeuGluLysSerHisCysIleAlaGluValGluAsnAspGluMetPro  
 CTGAGGAAATGCTGAAACCTCTGTTGGAAAAATCCACTGCATTGCCGAAAGTGGAAAATGATGAGATGCC  
 901 911 921 931 941 951 961 971  
  
 AlaAspLeuProSerLeuAlaAlaAspPheValGluSerLysAspValCysLysAsnTyrAlaGluAlaLysAsp  
 GCTGACTTGCCTTCATTAGCGGCTGATTTGTTGAAAGTAAAGGATTTGCAAAACTATGCTGAGGCAAGGAT  
 976 986 996 1006 1016 1026 1036 1046  
  
 ValPheLeuGlyMetPheLeuTyrGluTyrAlaArgArgHisProAspTyrSerValValLeuLeuLeuArgLeu  
 GTCTCTGGCAATGTTTGTATGAAATGCAAGAAGGATCCTGATTAATCTGCTGACTGCTGAGACTT  
 1051 1061 1071 1081 1091 1101 1111 1121

Figure 7

## Pl. VIII/31

AlaLysThrTyrGluThrThrLeuGluLysCysCysAlaAlaAlaAspProHisGluCysTyrAlaLysValPhe  
 GCCAAGACATATGAAACCCTCTAGAGAAGTGCTGCGCTGCAGATCCTCATGAATGCTATGCCAAAGTGTTC  
 1126 1136 1146 1156 1166 1176 1186 1196  
  
 AspGluPheLysProLeuValGluGluProGlnAsnLeuIleLysGlnAsnCysGluLeuPheGluGlnLeuGly  
 GATGAATTAAACCTCTGTGGAAGAGCCTCAGAATTAAATCAAACAAAATTGTGAGCTTTGAGCAGCTTGGAA  
 1201 1211 1221 1231 1241 1251 1261 1271  
  
 GluTyrLysPheGlnAsnAlaLeuLeuValArgTyrThrLysLysValProGlnValSerThrProThrLeuVal  
 GAGTACAATTCCAGAATGCGCTATTAGTTCTCACCAAGAAAGTACCCCAAGTGTCAACTCCAACCTTGTA  
 1276 1286 1296 1306 1316 1326 1336 1346  
  
 GluValSerArgAsnLeuGlyLysValGlySerLysCysCysAsnProGluAlaLysArgMetProCysAlaGlu  
 GAGGTCTCAAGAAACCTAGGAAAGTGGCAGCAAATGTTGTAATCCTGAAGCAGAAAGAATGCCCTGTGCAGAA  
 1351 1361 1371 1381 1391 1401 1411 1421  
  
 AspTyrLeuSerValValLeuAsnGlnLeuCysValLeuHisGluLysThrProValSerAspArgValThrLys  
 GACTATCTATCCGTGGCTCTGAACCAGTTATGTGTTGCATGAGAAAAGGCCAGTAAGTGACAGAGTCACCAA  
 1426 1436 1446 1456 1466 1476 1486 1496  
  
 CysCysThrGluSerLeuValAsnArgArgProCysPheSerAlaLeuGluValAspGluThrTyrValProLys  
 TGCTGCACAGAACATCCTGGTGAACAGGGCACCATGCTTTCTGAAGTCTGGAGTCAACATACGTTCCCAA  
 1501 1511 1521 1531 1541 1551 1561 1571  
  
 GluPheAsnAlaGluThrPheThrPheHisAlaAspIleCysThrLeuSerGluLysGluArgGlnIleLysLys  
 GAGTTAACGAACTTCACCTCAGCAGATAATGACACTTCTGAGAAGGAGAGACAAATCAAGAAA  
 1576 1586 1596 1606 1616 1626 1636 1646  
  
 GlnThrAlaLeuValGluLeuValLysHisLysProLysAlaThrLysGluGlnLeuLysAlaValMetAspAsp  
 CAAACTGCACCTTGTTGAGCTTGAAACACAAGCCCAAGGCAACAAAAGAGCAACTGAAAGCTGTTATGGATGAT  
 1651 1661 1671 1681 1691 1701 1711 1721  
  
 PheAlaAlaPheValGluLysCysCysLysAlaAspAspLysGluThrCysPheAlaGluGluGlyLysLysLeu  
 TTGCGAGCTTTGTAGAGAAGTGCTGCAAGGCTGACGATAAGGAAACCTGCTTGCCGAGGAGGTAAAAACTT  
 1726 1736 1746 1756 1766 1776 1786 1796  
  
 ValAlaAlaSerGlnAlaAlaLeuGlyLeu\*\*\*  
 GTTGCTGCAAGTCAAGCTGCCCTTAGGCTTATAACATCACATTAAAAGCATCTCAGCCTACCATGAGAATAAGAG  
 1801 1811 1821 1831 1841 1851 1861 1871  
  
 AAAGAAAATGAAGATCAAAAGCTT  
 1876 1886 1896

Figure 7

2635115

Pl. IX/31

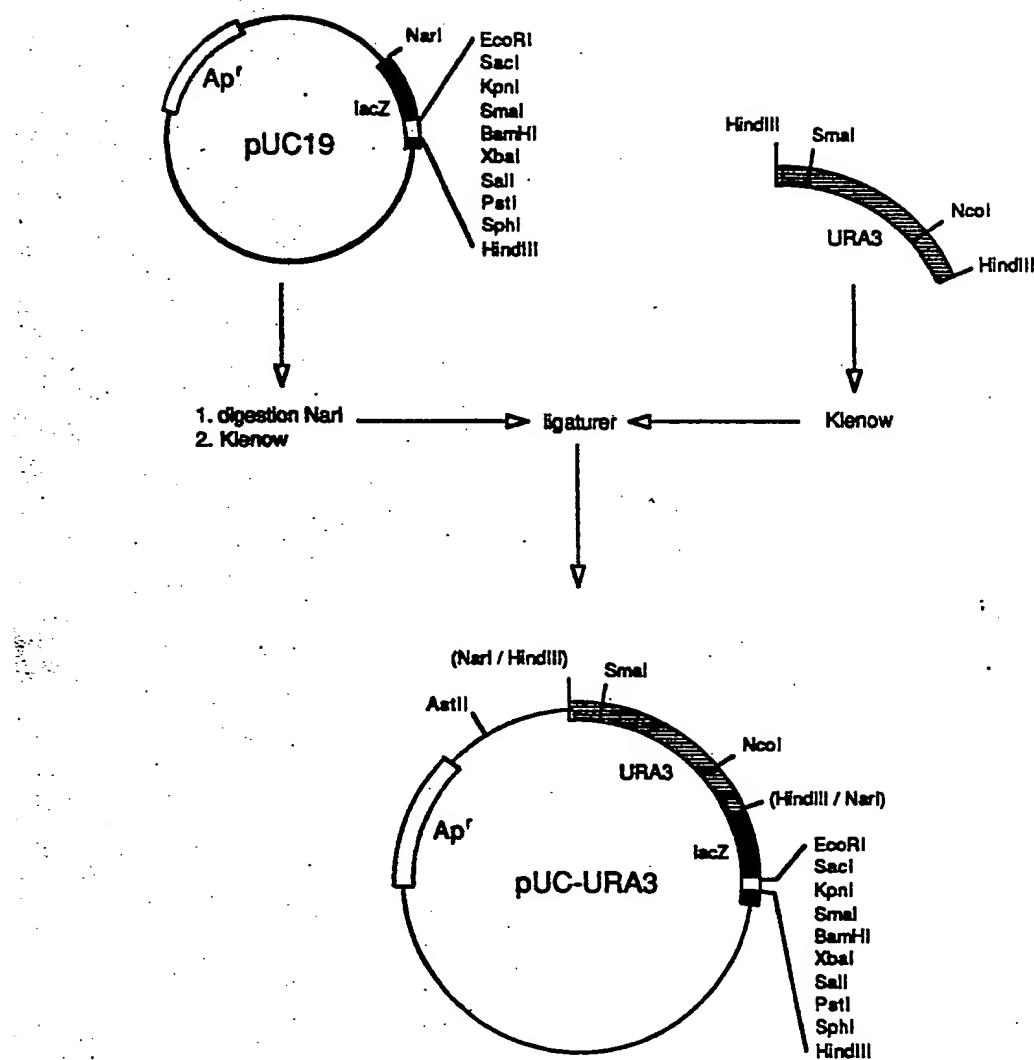


Figure 8

Pl. X/31

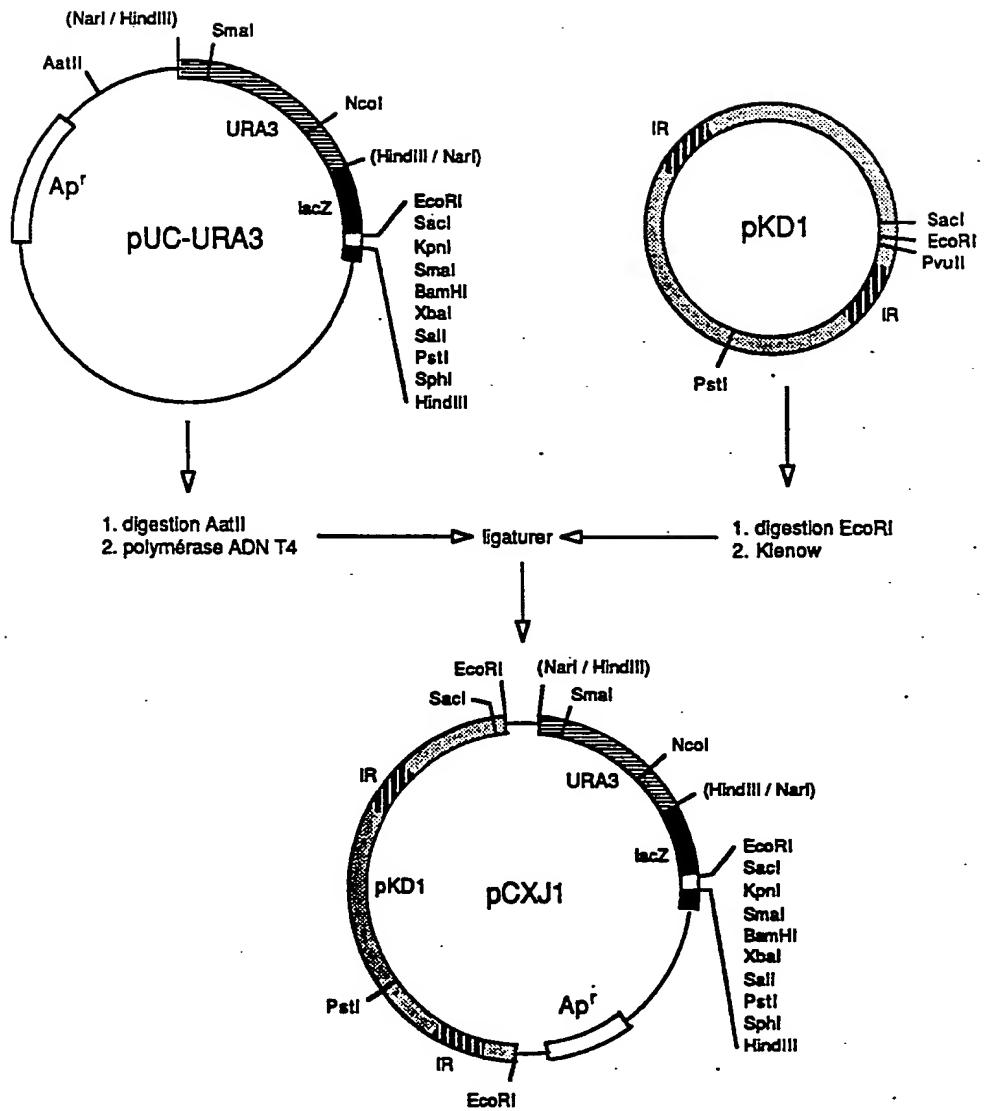


Figure 9

2635115

Pl XI/31

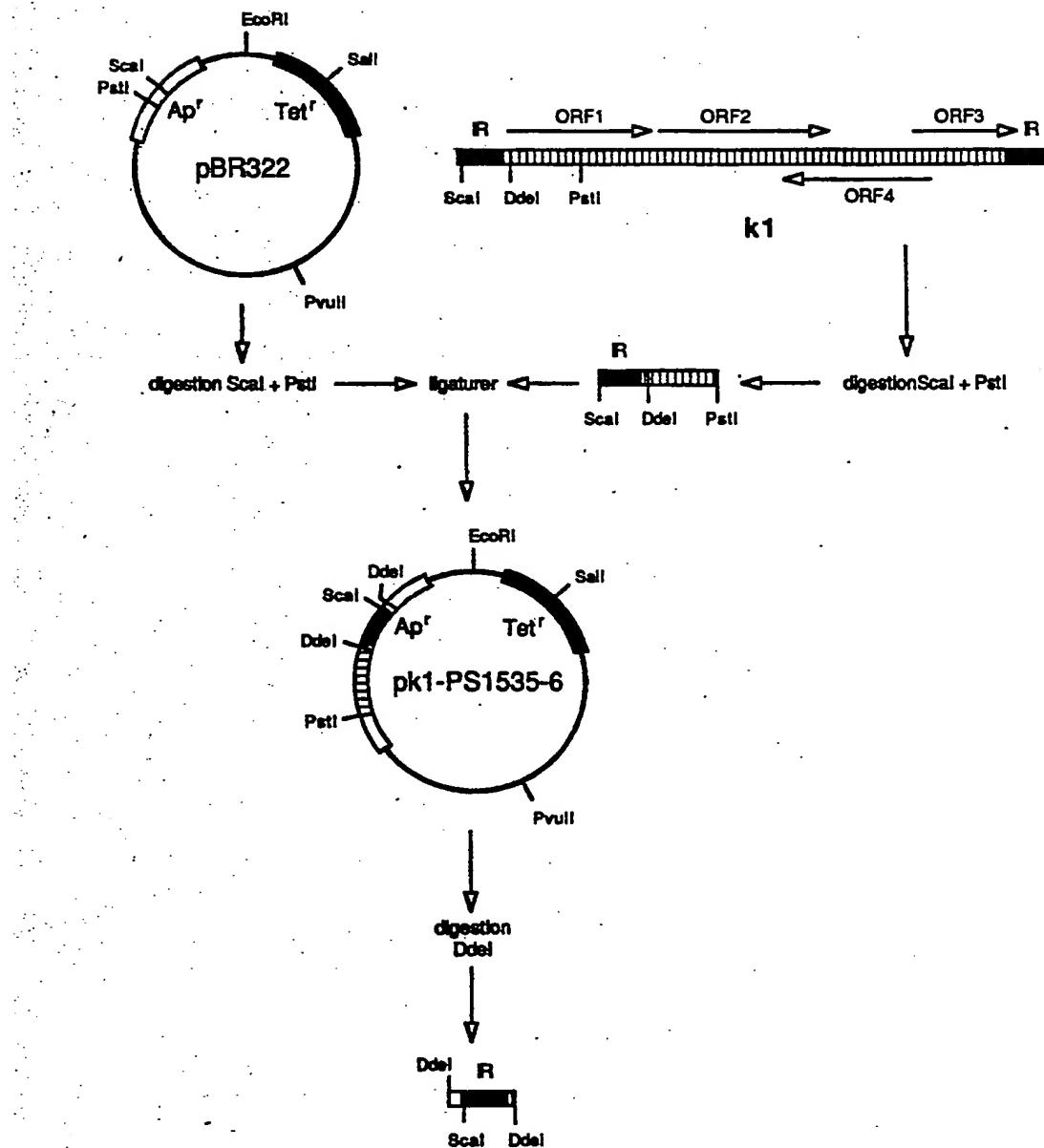


Figure 10

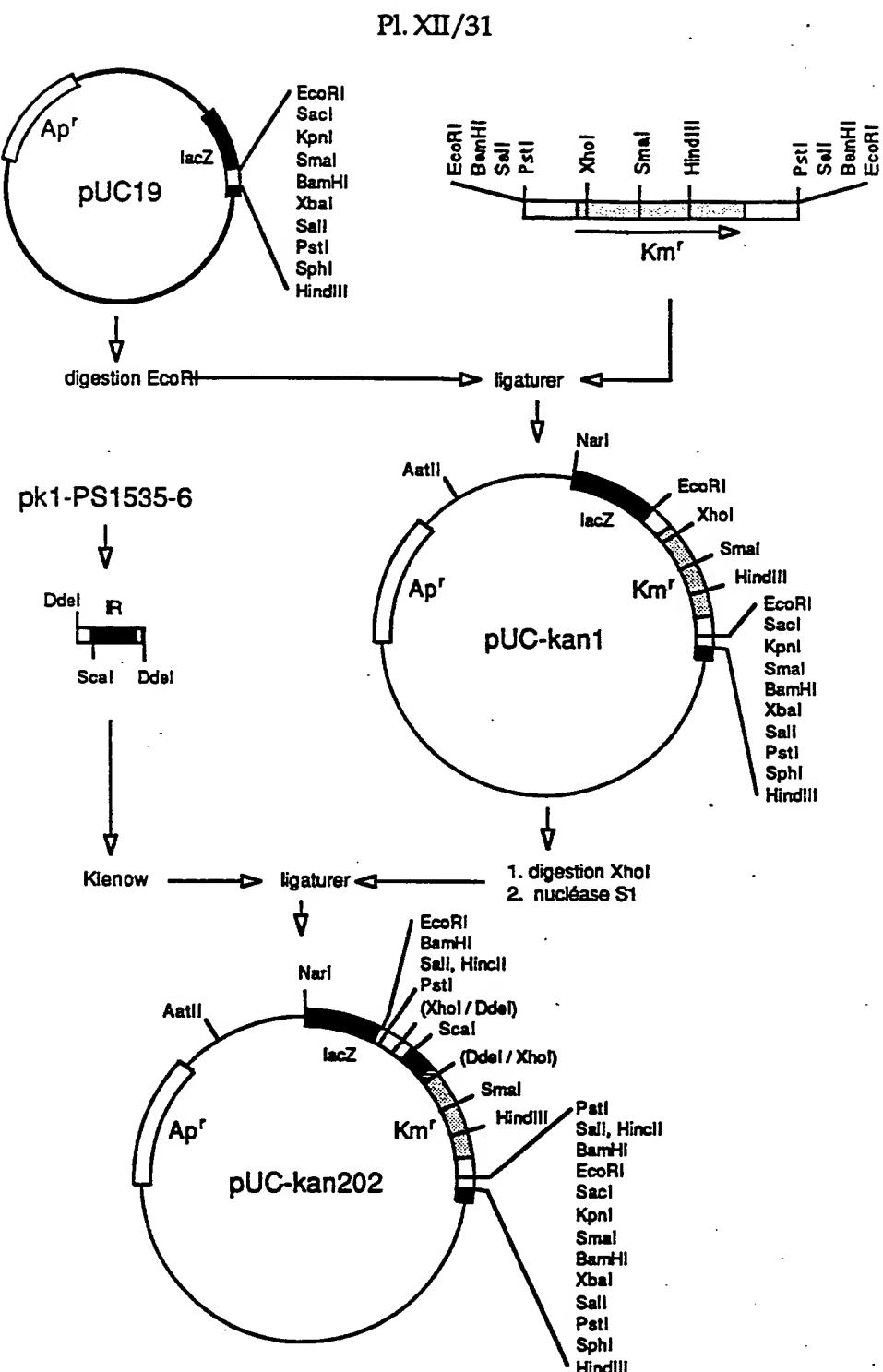


Figure 11

2635115

PL XIII/31

promoteur pk1

Scal

5' - AGTACTAAAAGTGAGATTATTGGAAGATTAGATCGTCTCCATTTCCTGTTTTGTTTAT

ATATTAGGTATTTTTTCAGTTTATATCTTCCTGTATAACAAGTCTATTTTTATATTTAAGTC

TATTTTACACTTATGACCTATAAGTCATTATTACACATTCCAACCTATAATATATGAATTACATT

ATTAATTTAAAA | ATG GAT TTC AAA GAT AAG GCT TTA AAT GAT CTA | AGG CCG CGA  
ORF1' (DdeI/XbaI) ORF1' ← → 'aph'

TTA AAT TCC AAC ... - 3'

Figure 12

## PL XIV/31

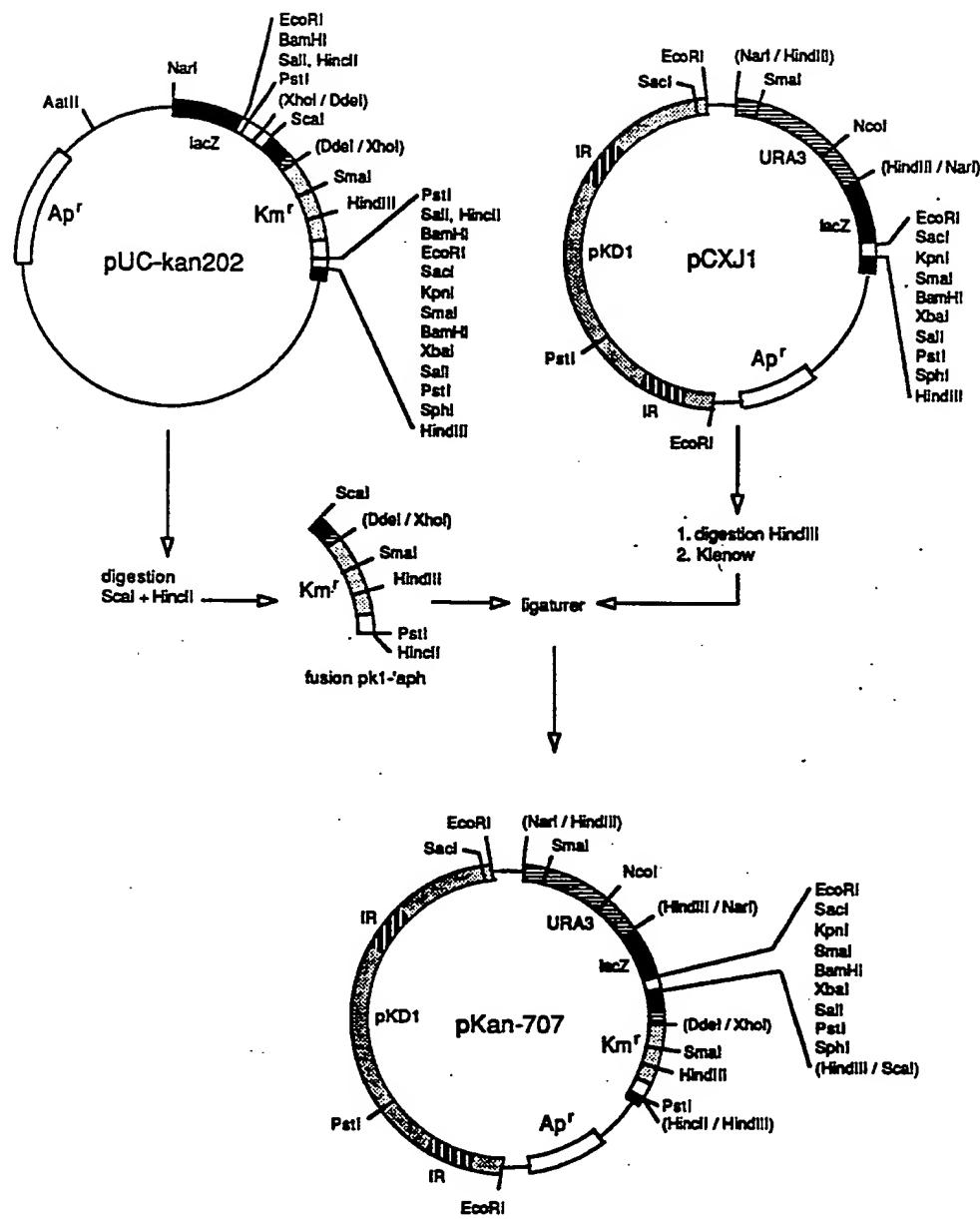


Figure 13

2635115

Pl XV/31

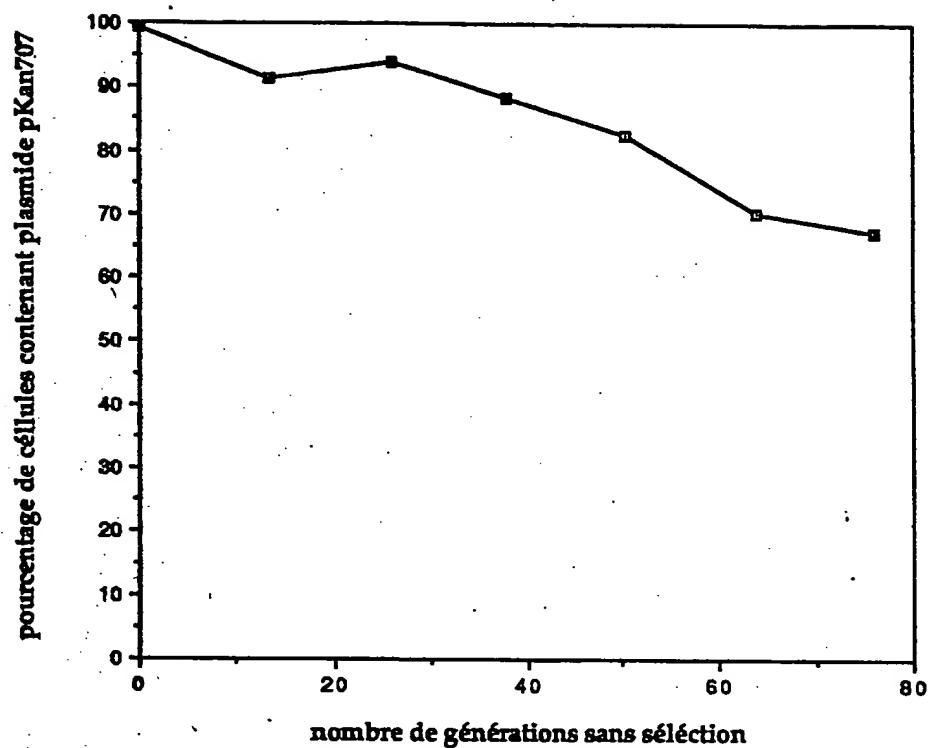


Figure 14

## Pl XVI/31

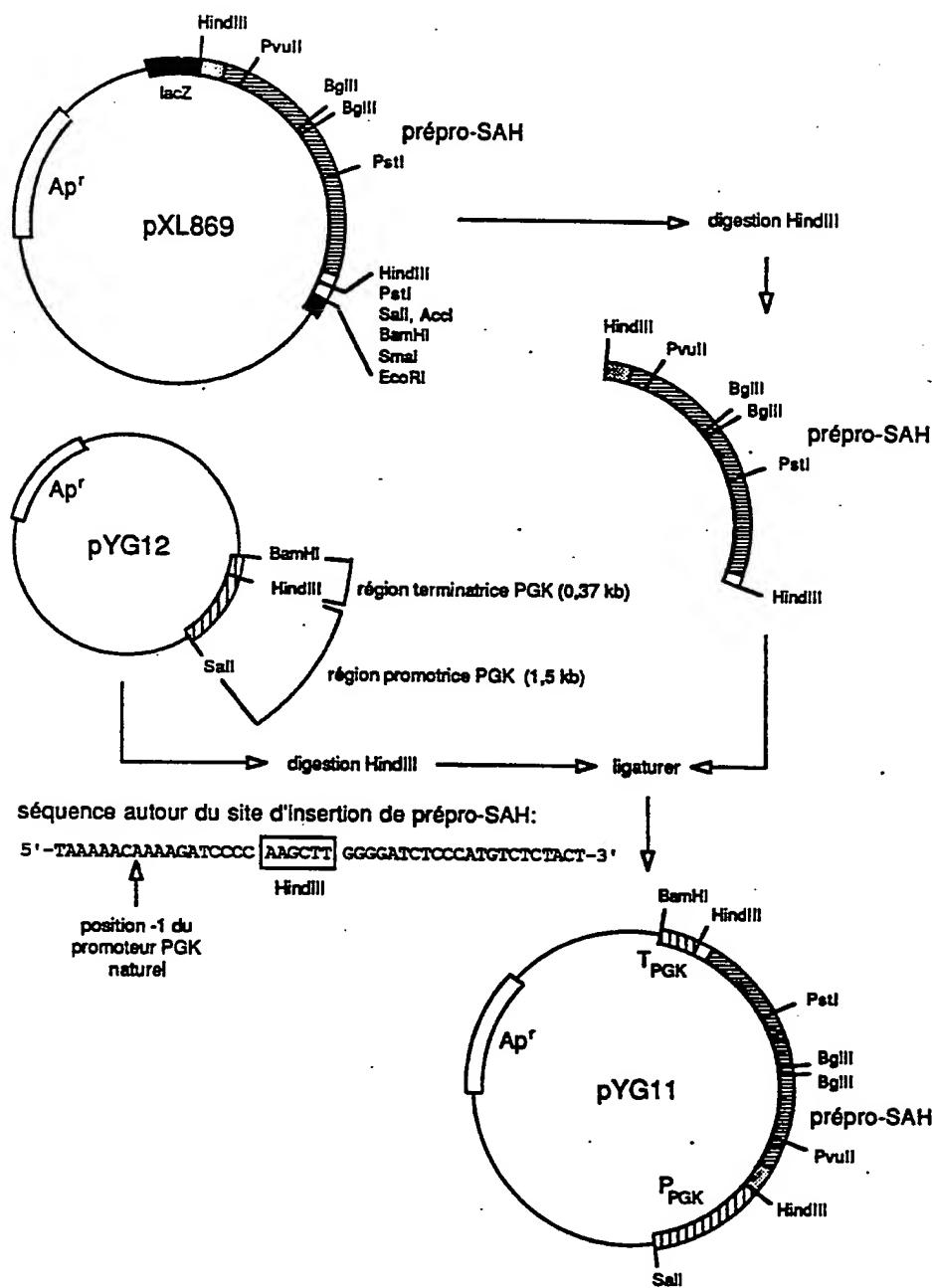


Figure 15

Pl. XVII/31

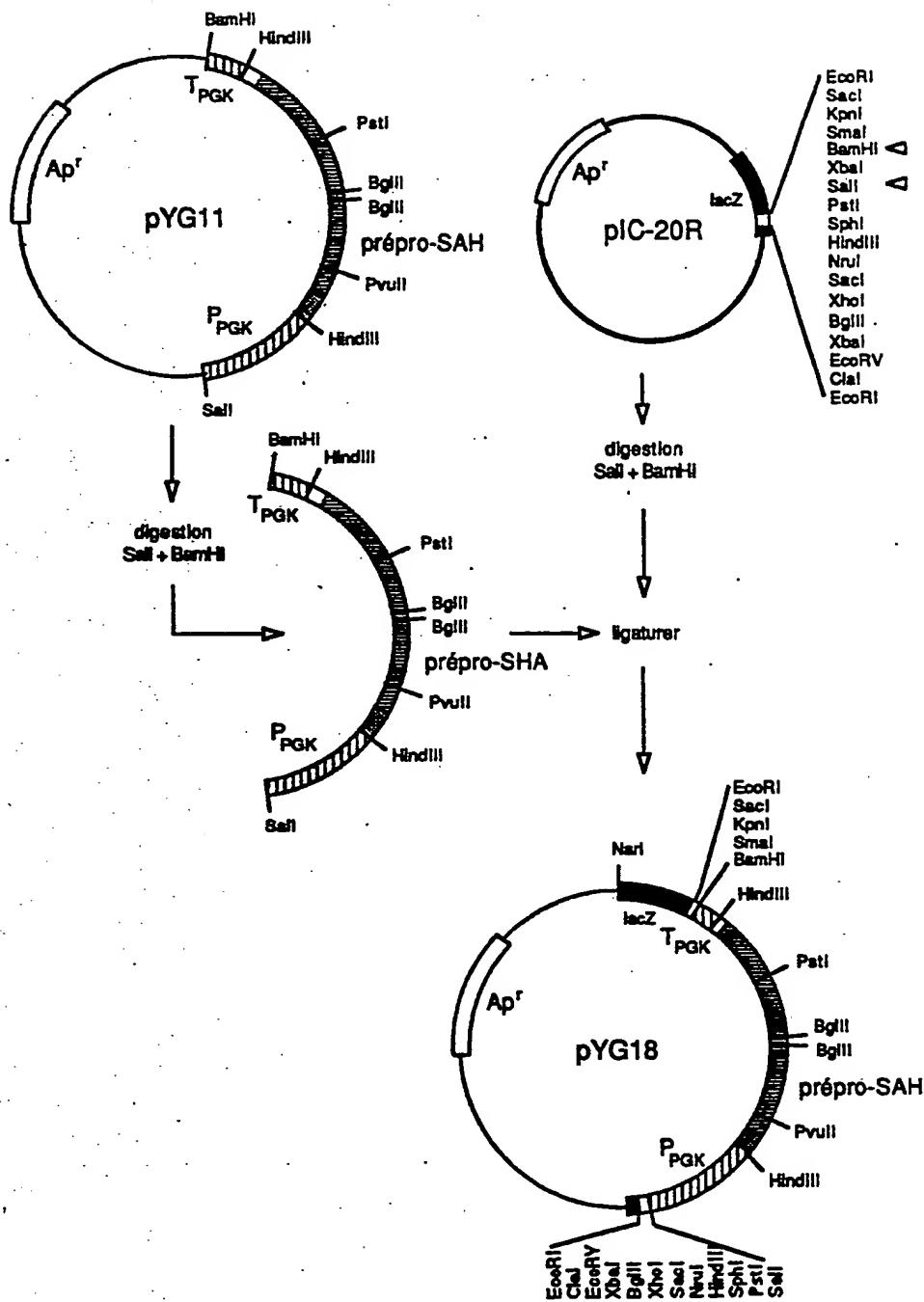


Figure 16

Pl. XVIII/31

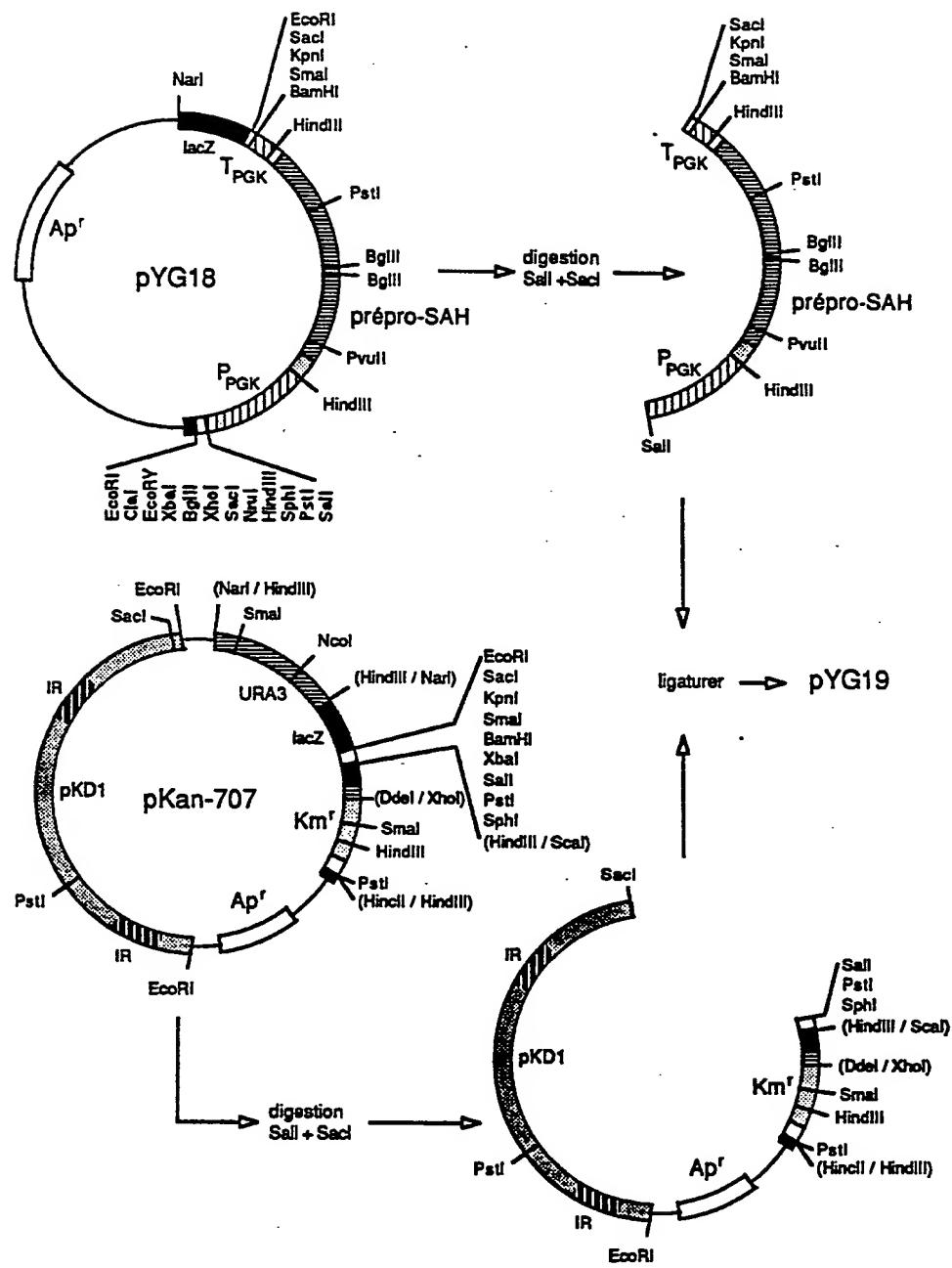
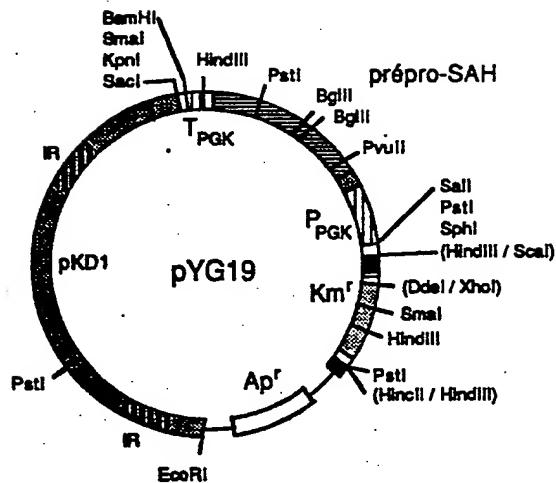


Figure 17

2635115

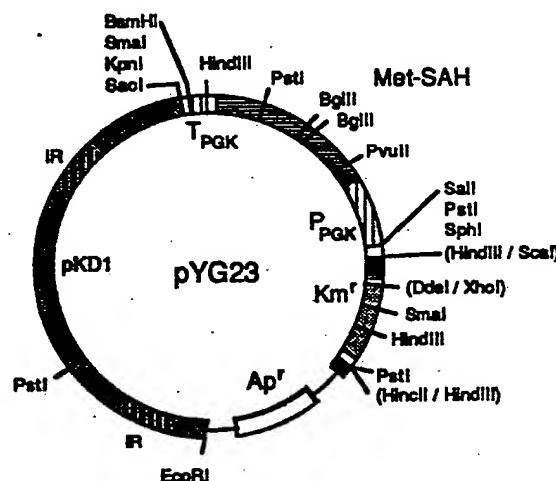
Pl. XIX/31



The diagram shows the SV40 early promoter region with three transcription start sites:

- pre**: Located at the top, starting from the first nucleotide (5' end).
- pro**: Located in the middle, starting after the first 10 nucleotides.
- SAH**: Located at the bottom, starting after the first 12 nucleotides.

The sequence is: 5' - TAAAGACA AAAGATCCCCAAGCTT ATG AAG TGG GTA ACC TTT ATT TCC CTT CTT TTT CTC ... - 3'



$\text{P}_{\text{PGK}}$   
 5' - TAAAAACA AAAGATCCCCAAGCTT ATG GAT GCA CAC AAG AGT ... - 3'

**Figure 18**

Pl. XX/31

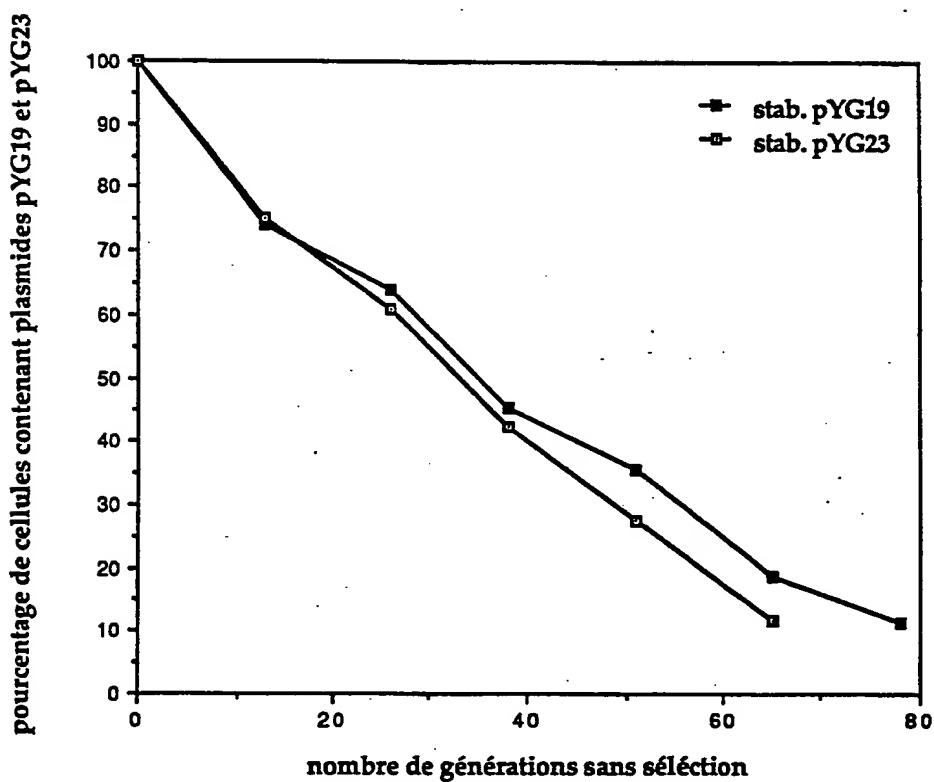


Figure 19

2635115

Pl. XXI/31

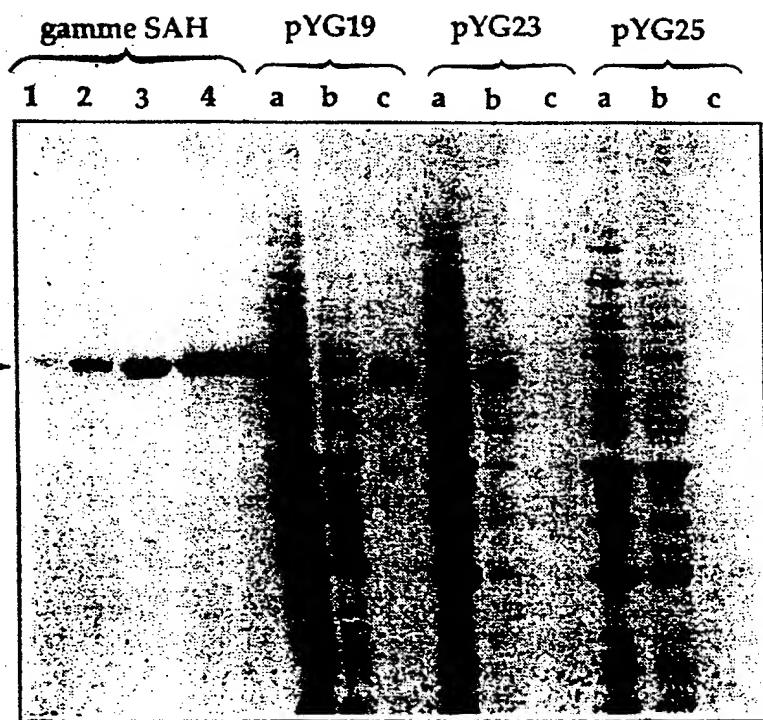


Figure 20

2635115

Pl. XXII/31

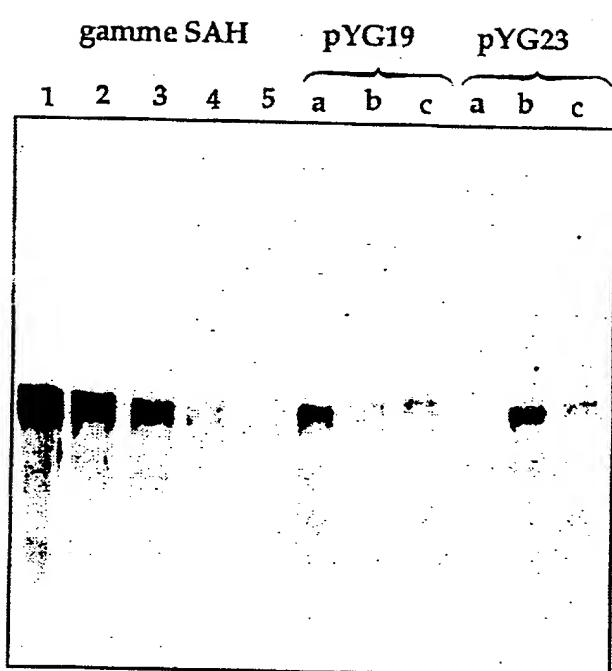


Figure 21

2635115

Pl. XXIII/31

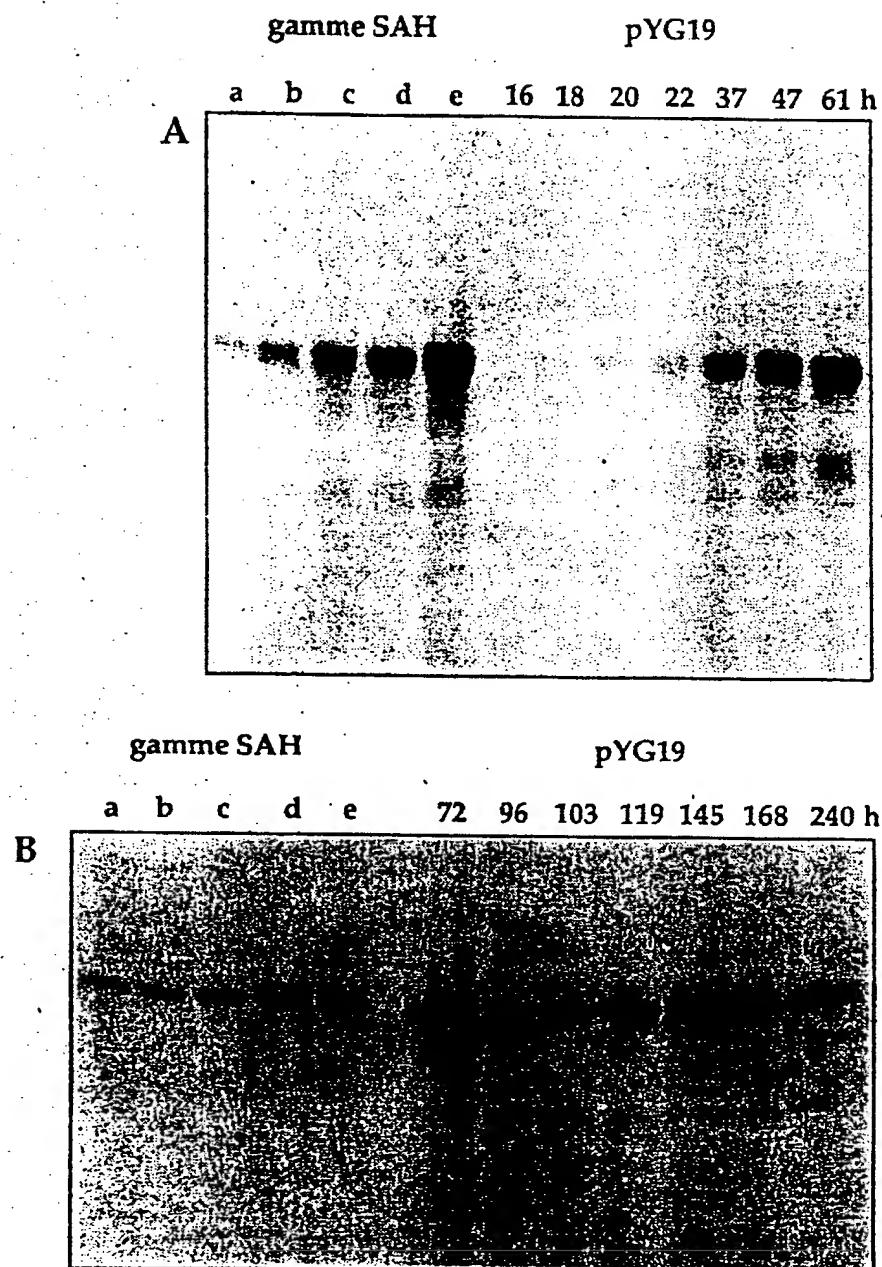


Figure 22

2635115

PL XXIV/31

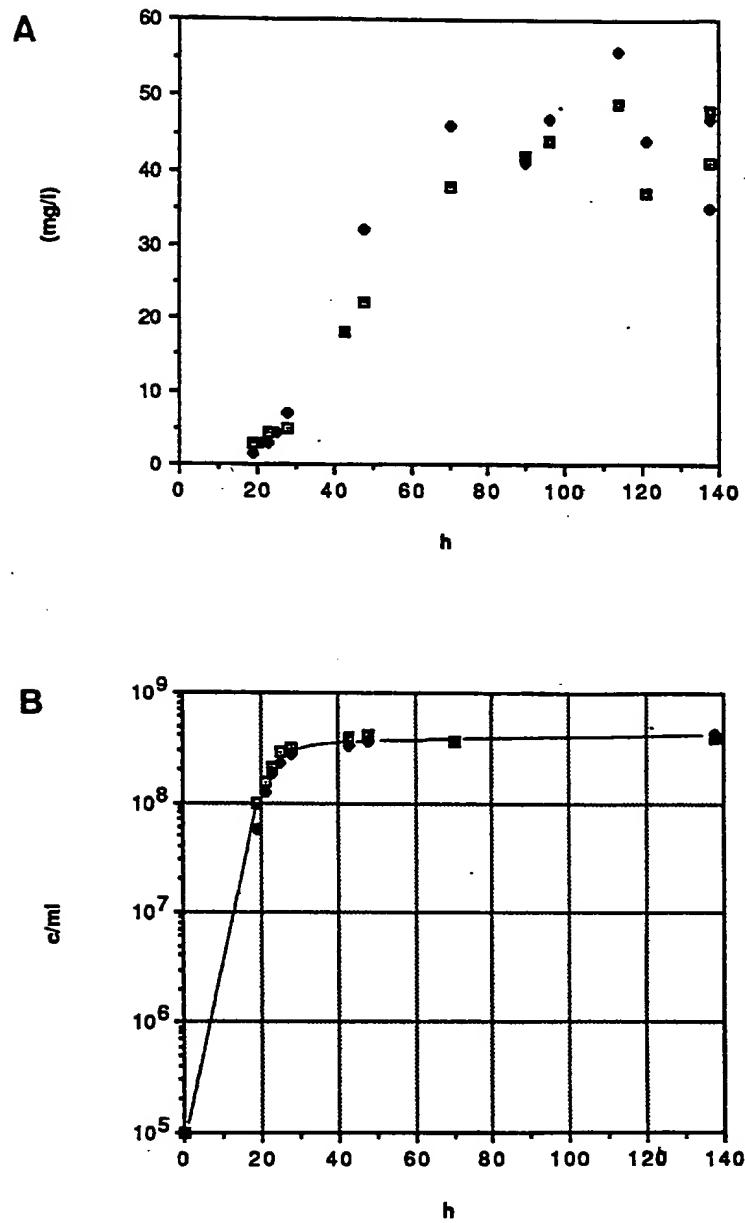


Figure 23

2635115

PL XXV/31

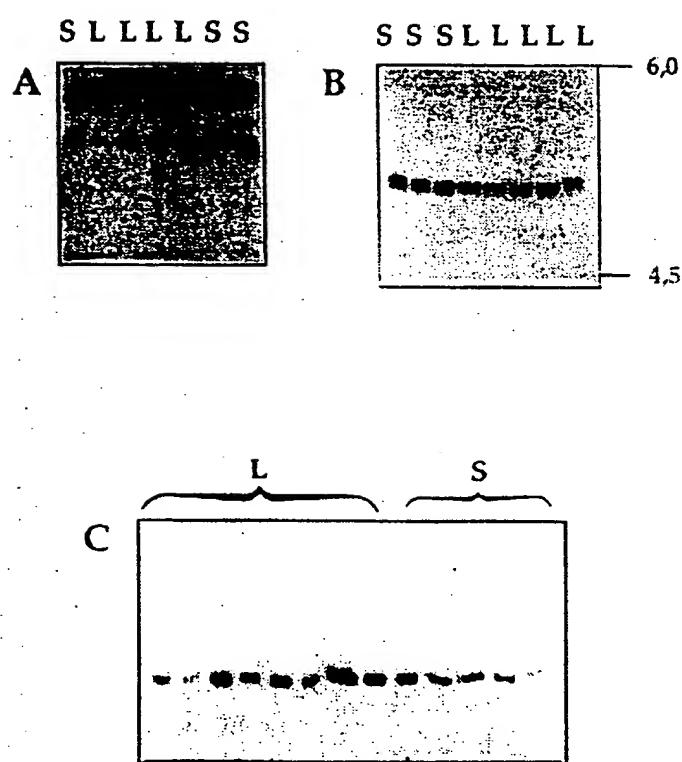


Figure 24

2635115

PL XXVI/31

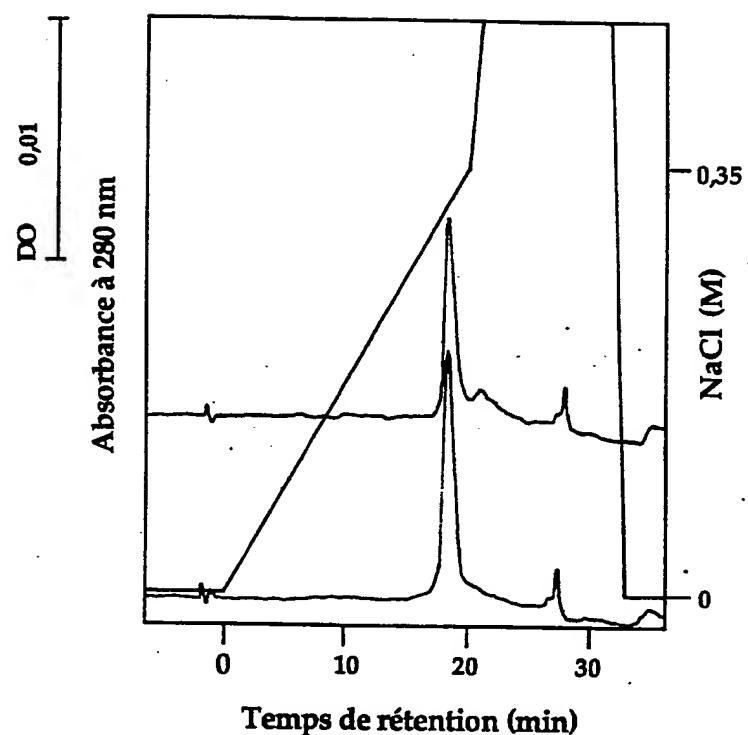


Figure 25

2635115

Pl. XXVII/31

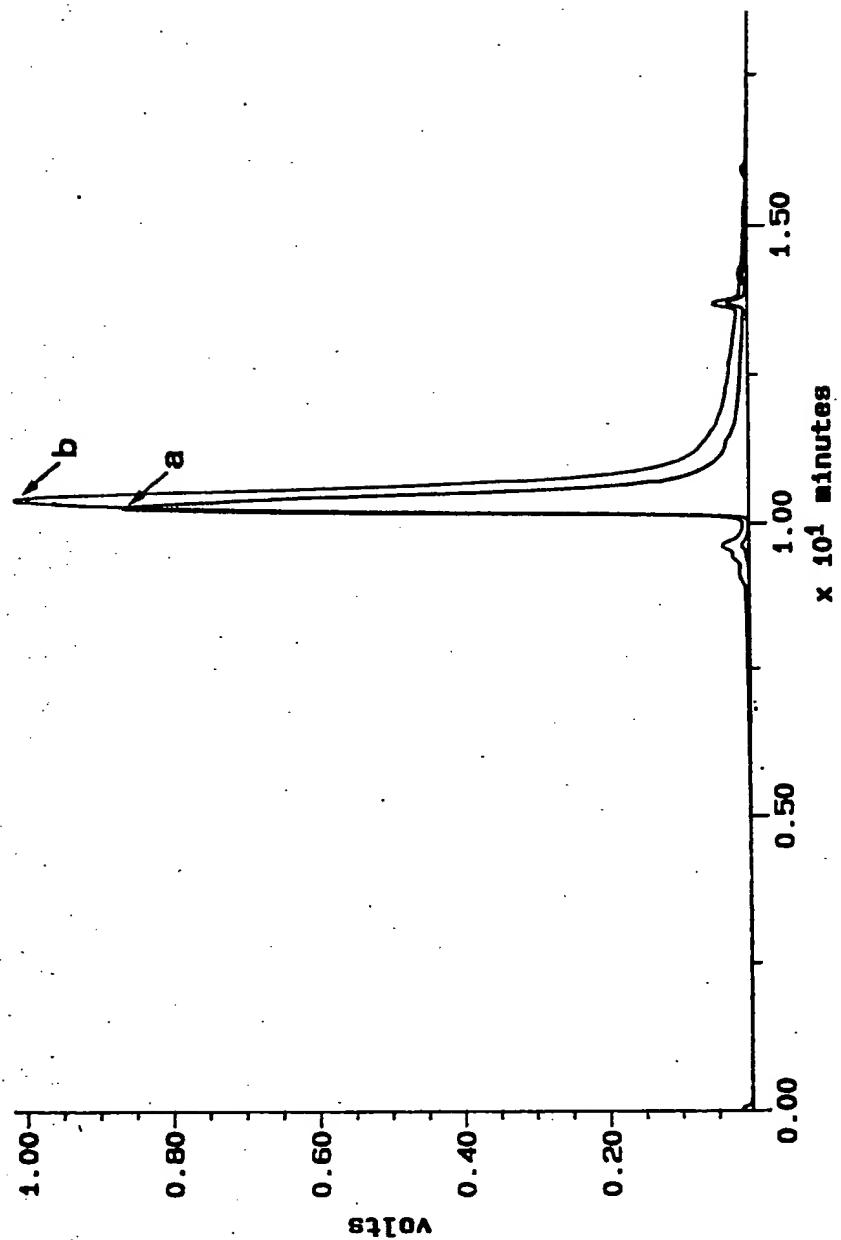


Figure 26

2635115

PL XXVIII/31

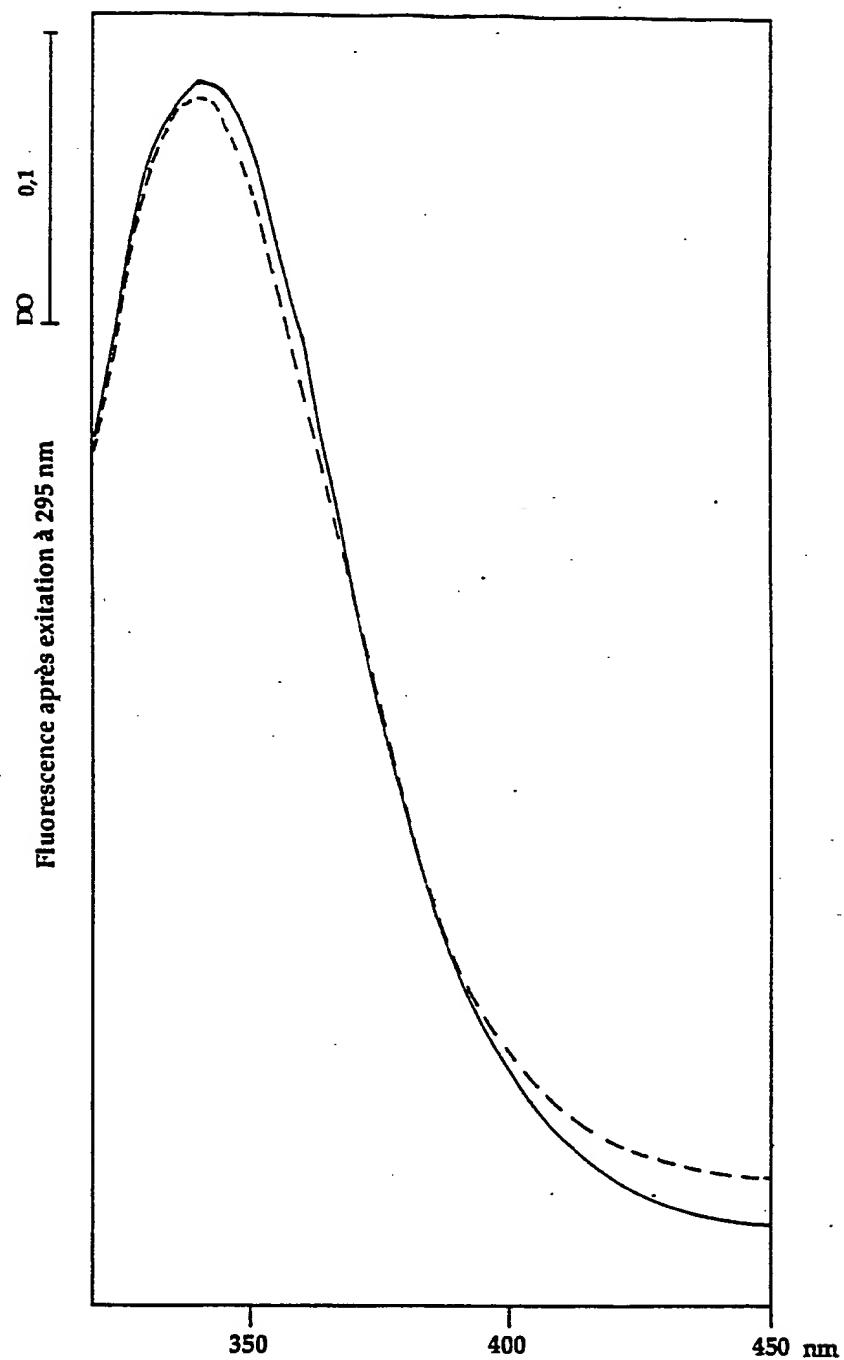


Figure 27

2635115

Pl. XXIX/31

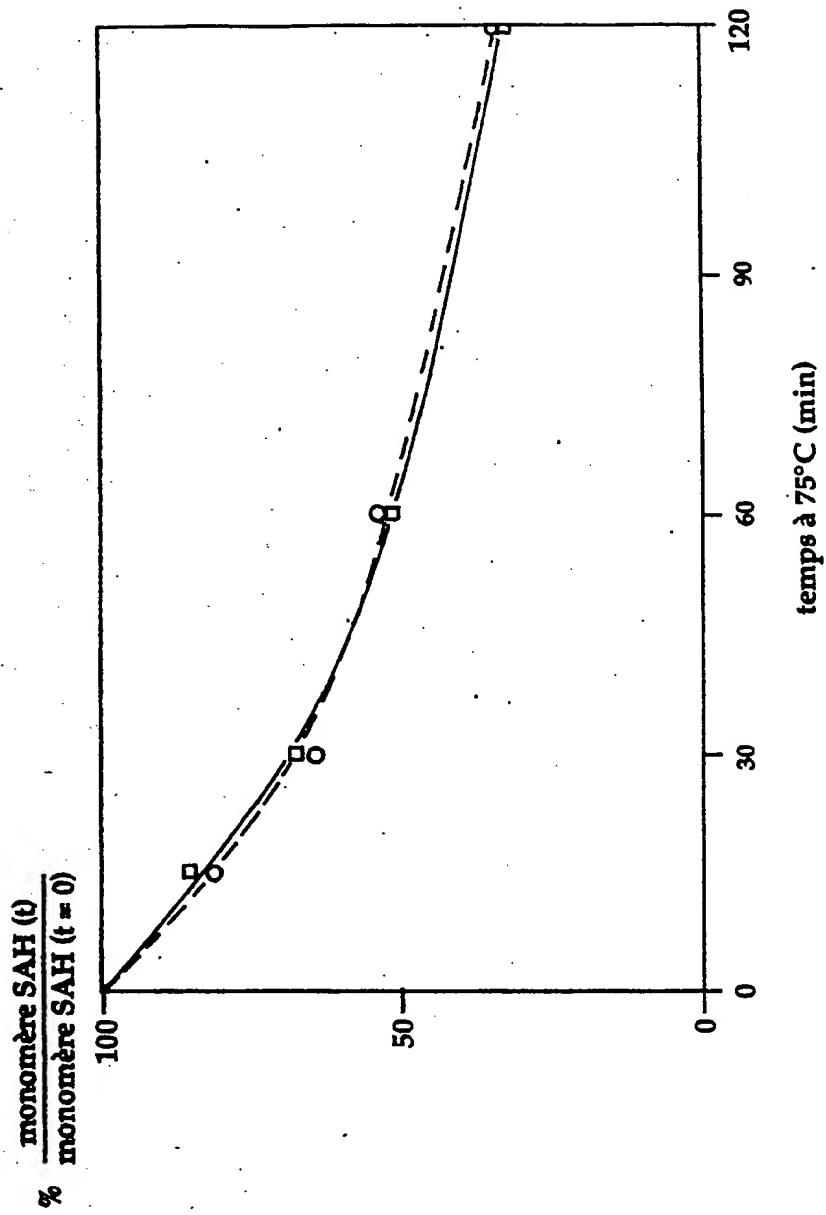


Figure 28

2635115

Pl. XXX/31

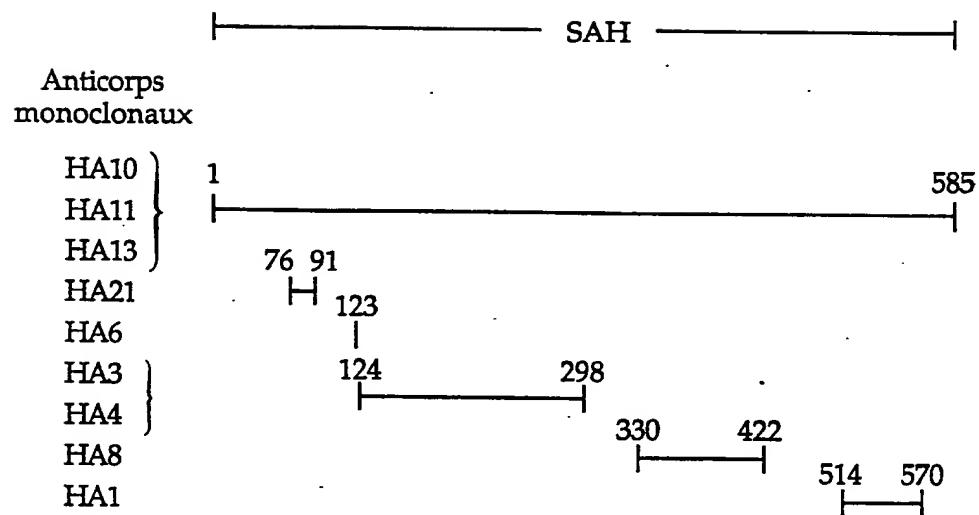


Figure 29

2635115

Pl. XXXI/31

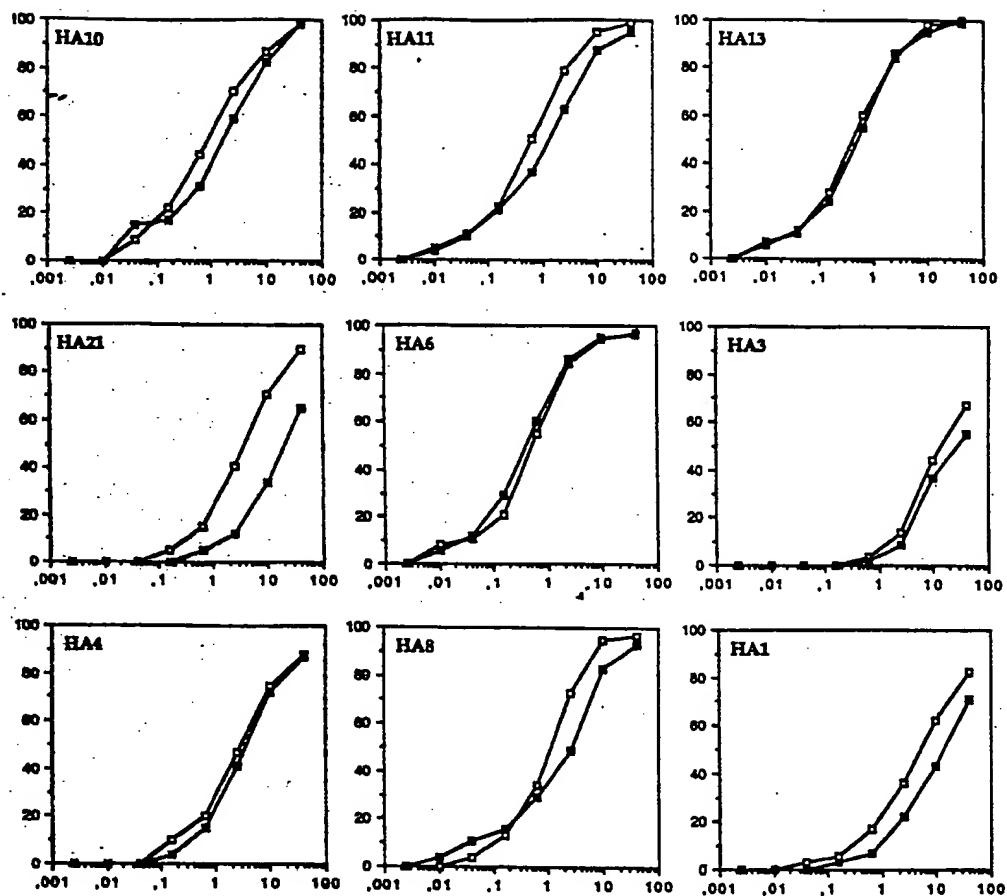


Figure 30